

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 38.5 Seconds

(without alignments)
4564.770 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLLTGALGLAFAFASWAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mbc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	425	14.9	576	Q8J2Q5	Q8J2Q5 gibberella
2	419	14.7	593	10 Q9C6S0	Q9C6S0 arabidopsis
3	419	14.7	595	10 Q93VK5	Q93VK5 arabidopsis
4	417	14.6	503	6 Q97689	Q97689 sus scrofa
5	417	14.6	517	10 Q8LIR5	Q8LIR5 oryza sativ
6	403.5	14.1	503	6 Q95L3S	Q95L3S sus scrofa
7	403	14.1	515	10 Q9LUD2	Q9LUD2 arabidopsis
8	401	14.1	515	10 Q8WIE1	Q8WIE1 arabidopsis
9	400.5	14.0	519	10 Q9ASR3	Q9ASR3 arabidopsis
10	397.5	13.9	519	10 Q8LHV0	Q8LHV0 oryza sativ
11	394	13.8	503	4 Q86SK3	Q86SK3 homo sapien
12	393.5	13.8	544	10 Q9F3E1	Q9F3E1 oryza sativ
13	392	13.7	497	11 Q8CJF2	Q8CJF2 rattus norv
14	389	13.6	492	5 Q8IS07	Q8IS07 mamestra br
15	387	13.6	523	10 Q8LQ77	Q8LQ77 oryza sativ
16	386.5	13.6	511	6 Q8HYV5	Q8HYV5 capra hircu

17	385.5	13.5	535	4 Q9H241	Q9H241 homo sapien
18	384.5	13.5	531	10 Q9FED3	Q9FED3 oryza sativ
19	382.5	13.4	527	10 Q8LGM8	Q8LGM8 zea mays su
20	382	13.4	430	4 Q7Z448	Q7Z448 homo sapien
21	381	13.4	503	6 Q8HZK1	Q8HZK1 canis famil
22	381	13.4	520	10 Q8LQ04	Q8LQ04 oryza sativ
23	380.5	13.3	505	10 Q9LUD0	Q9LUD0 arabidopsis
24	374	13.1	533	5 Q964T2	Q964T2 blattella g
25	371	13.0	507	10 Q8LQ08	Q8LQ08 oryza sativ
26	368.5	12.9	523	10 Q9FD21	Q9FD21 oryza sativ
27	365	12.8	512	10 Q9LUC8	Q9LUC8 arabidopsis
28	365	12.8	520	10 Q48786	Q48786 arabidopsis
29	364.5	12.8	503	5 Q8WQ89	Q8WQ89 anopheles g
30	363.5	12.7	511	10 Q64631	Q64631 arabidopsis
31	363.5	12.7	512	10 Q9LUD3	Q9LUD3 arabidopsis
32	363	12.7	502	11 Q06884	Q06884 rattus norv
33	362.5	12.7	523	10 Q9CAD6	Q9CAD6 arabidopsis
34	362	12.7	506	10 Q9LUD1	Q9LUD1 arabidopsis
35	361	12.7	528	10 Q9FE60	Q9FE60 oryza sativ
36	360	12.6	512	10 Q8W2N3	Q8W2N3 vicia sativ
37	359.5	12.6	512	10 Q9LUC9	Q9LUC9 arabidopsis
38	357.5	12.5	512	10 Q9LUC5	Q9LUC5 arabidopsis
39	357.5	12.5	528	10 Q8LL74	Q8LL74 zea mays su
40	356.5	12.5	523	10 Q8LAQ9	Q8LAQ9 arabidopsis
41	356	12.5	504	11 Q9EQW4	Q9EQW4 mus musculu
42	355.5	12.5	512	10 Q9ZW95	Q9ZW95 arabidopsis
43	354.5	12.4	506	10 Q7XTY4	Q7XTY4 oryza sativ
44	354.5	12.4	560	10 Q8W372	Q8W372 oryza sativ
45	354.5	12.4	560	10 Q7XD91	Q7XD91 oryza sativ

ALIGNMENTS

RESULT 1
Q8J2Q5 PRELIMINARY; PRT; 576 AA.
ID Q8J2Q5
AC Q8J2Q5; 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Fum15p.
GN FUM15.
OS Gibberella moniliformis.
OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
CX NCBI_TaxID=1171187;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M-3125;
RX MEDLINE=99343882; PubMed=10413619;
RA Proctor R.H., Desjardins A.E., Plattner R.D., Hohn T.M.;
RT "A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in Gibberella fujikuroi mating population A.";
RL Fungal Genet. Biol. 27:100-112(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M-3125;
RX MEDLINE=21585628; PubMed=11728154;
RA Seo J.A., Proctor R.H., Plattner R.D.;
RT "Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in Fusarium verticillioides.";
RL Fungal Genet. Biol. 34:155-165(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=M-3125;
RX Proctor R.H., Brown D.W., Plattner R.D., Desjardins A.E.;
RT "Co-expression of fifteen contiguous genes delineates a fumonisin biosynthetic gene cluster in Gibberella moniliformis.";
RL Fungal Genet. Biol. 0:0-0(2003).
DR EMEL; AF155773; AAN74818.1; -;
DR GO; GO:0003676; F:nicotinic acid binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Atg31800/68069.m00159.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RA "Arabidopsis cdna clones";
 RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RA "Arabidopsis cdna clones";
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 RA Huan W.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
 RA Yamada K., Yu G., Yuan S., Davis R.W., Theologis A., Ecker J.R.;
 RA "Arabidopsis ORF clones";
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AY058173; AAL25587.1; -;
 DR EMBL; AY056445; AAL08302.1; -;
 DR EMBL; AY142017; AAM98281.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 595 AA; 66845 MW; A548A33A48232B7D CRC64;

QY 206 ALAQFYKSLQNTKNTLYVAVGLTDPAPTLDSFKAINMDVFPYFTWTKRHEIPLT 265
 Db 256 GKAVFNDFSLNDTGVIEAVTVREADRSPSP--VMD-IPWKDIS-----PRQ 307
 QY 266 QGLAVRRVGIEMEQKQAVLGASDAQVDKDVQ-----ORD--ILSLVANIAA 316
 Db 308 RKVATSLKLNIDTLD-----LIATCKMVEEELQFHEEYMERDPSILHFL----- 355
 QY 317 NLPSQKLSDEEVLQAIISLLFAGVETSTVLTWTFHRLSEDKAVOKLREECIQDID- 375
 Db 356 -LASGDDVSKQLRDDIMTMTLAGHETSAAVLTWTFYLLTTPESVAVKQEEVDVIGDR 414
 QY 376 MPTLDELNALPYLEAVKESLMDPPSPYANRECKNDEDFIPLAEPVIGDGSVINEVRI 435
 Db 415 FFIQDQKKLKYTRVMNESLRYPPVPLRRSI-----ENDILGYPYI 459
 QY 436 TKGTMVMLPLFINRKSFTYGEDAEFRPERWLEVDVD-----SLNSIEAPYHQASPIG 491
 Db 460 KRGEIDFISVWNLHRSPLHWDADAEKFNPERWPLDGNPNETNQNFSLPFG-----GG 512
 QY 492 PRACFGWRPFAVEMKAFVLTLLRRVQFE 519
 Db 513 PRKIGDMFASFENVVAIAMLIRRFNFQ 540
 RESULT 4
 O97689 PRELIMINARY; PRT; 503 AA.
 ID O97689
 AC O97689
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome P450.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schuetz E., Andag R., Wieland E., Oellerich M.;
 RT "Porcine Cytochrome-P450 with high protein homology to human CYP3A4
 RT (Nifedipine oxidase J04449).";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF109068; AAD04628.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. . .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01689; EP450IICYP3A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 503 AA; 57267 MW; 22DB19F43C5988BB CRC64;

Query Match 14.8%; Score 419; DB 10; Length 595;
 Best Local Similarity 26.4%; Pred. No. 4.1e-22;
 Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NFDILSARTG--BEHAKYREKYGST-----LRPAGI-----AGAPVLN 86
 Db 95 NVLDFMDFMTGSDQDPKVPKANGSQAVNEAFFPLVELTYGIFRLTFGPKSFLI 154
 QY 87 STDPKVFNVHVK--AYDPKPGVAARVLRIATGDGVVTAEGEAHGRHRIMPSLSQAV 145
 Db 155 VSDPSIAKHILDKNAKYSK-GILAEITLDFVNGKGLIPADGIMRRRRRAIVPALHQYV 213
 QY 146 KSMVPIFLKGMELVDKMDAAEKOMAVGESAGEKKARLETGVDVVDKDWGRTLDVM 205
 Db 214 AAMISLFGESDRLCQKL--DAA-----ALKGEVEVESLFSRLTLDII 255

QY 175 GESAGEKATRLTEGVDVVDKWDYGRATLDVWALAGDYKSDSL-----QNKTNELV 226
 Db 168 -----KPTMTKIDIFGAYSMVDVITSTAGVNTDFLNNPQDPFFVENSCKLLKF 213
 QY 227 AFVGLTDFGAPTLDSEKAIWMDVFPVPTMKRRHEIPLTQGLAVS-----RRVGIELMEQKK 283
 Db 214 SF-----FSLI-----FPEL-----TPILEVLNVTLPKSVVNFMSIK 254
 QY 284 QAVLGASDQAVDKVQGRDILSLVRAVNIANLPSQKLSDEEVLQIQLNLLFAGYET 343
 Db 255 R-----NKESRLKDKQTHRVDFDLQMLNSQNSKETDTHKGLSDEELVAQGVFFIAGYET 309
 QY 344 SSTVLTWFMFHELSEKAVQKLEETCOI--DFTDMETLDELNALPYLFAFKESIRLDP 401
 Db 310 TSSLSLVLVELATHPDVQKLEEDATFPKALPSYDALQMEYLDNVNEILRLPYI 369
 QY 402 SPYANRECLDDEFIPLAEFVIGRDSVINEVRITKTMVLMPLFNINRSKFIYGEDAAE 461
 Db 370 AARLERVCKKQVE-----IHGVSVPKGTVMVVPFSIHRDPDLWPE-PEE 413
 QY 462 FPERWLEEDVTDLSNLEAEVGHQAFISQPRACFGWRFAVEMKAFVTLRLRVQREPI 521
 Db 414 FPERFSKKNKDSIN-----PITY-LFPGTGPRNCIGMRPALNMKALVRLVQNPSPKPC 468
 QY 522 -----ISHPEYEHITLIISPRIVG 541
 Db 469 KETQPLKLSQGLIQEKBILLKVVPRDGTVG 501
 RESULT 5
 ID QBLIR5 PRELIMINARY; PRT; 517 AA.
 AC QBLIR5
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 GN OJ1332.C12.12.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 clone:OJ1332.C12.12";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AP003752; BAC10039.1; -;
 DR Gramine; QBLIR5; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; Cytochrome_P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 517 AA; 58038 MW; 3C3E35745026E138 CRC64;
 Query Match 14.68; Score 417; DB 10; Length 517;
 Best Local Similarity 26.58; Pred. No. 4.6e-22;
 Matches 154; Conservative 95; Mismatches 202; Indels 130; Gaps 24;
 QY 6 LITGALGAAPSWASIAFFSLYIAP-----RRSSLYNQPNHNTFTGNFLDLSARTGE 61
 Db 7 MVAANAASVAFDAVAVKLVNRPRAITRLRAQGVGGPV-RFFSGMLGEI--RLRD 63
 QY 62 EHA-----KYREKYSTLRPAGIAGAPVLNSTDPKPVNHWMEAY 101
 Db 64 EGAGVLDVSHDVPVLPVQPHFRKRWIPLYGKTPNY-WFGARPTICLADVSMVRQVLSRT 122

QY 102 D-YPK--PGMAARVRLIATGDGVTAEGEAKRHRRIIMPISLSAQAQVKSVMPIELEKGM 157
 Db 123 GYFPTQSNPFFARLL-----GKGLVLTGDGEWKSKRVVHPAFNMDKLQMT----- 170
 QY 158 ELVDKMMEDAAEKMDMAGVBSAGEKKATRLTEGVDVQKMWGR-----AILDVMAAGF--D 211
 Db 171 ----VTMSDCAQMSISEWES-----ELGTGDIIVEIELSRFRFELTADVISHTAFSS 219
 QY 212 YKSD-----SLQNKTNELYVAFVGLTDFGAPTLDSEKAIWMDVFPVPTMKRRHEIP 263
 Db 220 YKEGQVFLAQRELO-----FLAF-----STFLSI-----QIP 247
 QY 264 LTQGLAVRRVGIELMOKKQAVLGASDAQVDDKDV--QGRDILSLVRAVNIANLPS 321
 Db 248 GSSLPYTKKLNKLTWSVDKVRSMLTDTIISRLNNKDVAGYGNLGLMLEA-CAPEHGES 306
 QY 322 Q-KLSDSEVLQAISNLLFAGYETSTVLTWFMFHLSEDKAVQDKLREETC-QIDTDMPTL 379
 Db 307 QPOLSMDEIIAECKTFFFGAGHDTTSHLLTWTWFLSTHPEWQEKLEEVATECDGKVPTG 366
 QY 380 DELNALPYLFAFKESIRLDPSPYANRECLDDEFIPLAEFVIGRDSVINEVRITKGT 439
 Db 367 DMLAKLKLVMNFLTETLRLYGPVAFIQRVNALE-----LGGITVPEGT 411
 QY 440 VMMLPLFNINRSKFIYGEDAAEERPERWLEEDVTDLSNLEAEVGHQAFISQPRACFGWR 499
 Db 412 VLSIPIATIHRRKEVWGEDADIFKPERFKNQVSKAGKYPNA---LLSFSSSPRACIGCN 467
 QY 500 FAVAEMKAFVTLRLRVQREPIISHPEYEHITLIISPR 538
 Db 466 FAMIEAKAVIAMILQRFSEF---TLSPKYVHVPTDVTILPEK 505
 RESULT 6
 ID Q95L35 PRELIMINARY; PRT; 503 AA.
 AC Q95L35
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Cytochrome P450 3A.
 GN CYP3A.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Soucek P., Zuber R., Anzenbacherova E., Linka M., Anzenbacher P.;
 RT "CDNA of minipig cytochrome P450 3A";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF242780; AAL13316.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:oxidoreductase activity, acting on paired d. . .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR008072; EP450_CYP3A.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01689; EP450IICYP3A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 503 AA; 57085 MW; 5A17128CA9E50914 CRC64;
 Query Match 14.14; Score 403.5; DB 6; Length 503;
 Best Local Similarity 27.88; Pred. No. 4.4e-21;
 Matches 152; Conservative 81; Mismatches 191; Indels 123; Gaps 22;
 QY 12 GLAAPSASTA--PFSLYLAPRS-----SLYLNQPNHNTFTGNFLDLSARTGEH-- 63
 Db 6 GFSTETWLLATSLVLDLYLYGTYSHGFLFKLIGIPGRPLPYF-GN---ILGVRKGVDFD 61


```

OX NCBI_TaxID=3702;
RP (1)
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RP (2)
RP SEQUENCE FROM N.A.
RC Rounslev S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RC Strain=sv. Columbia;
RC STRAIN=sv. Columbia;
RA Shen W., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RP Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RP (3)
RP SEQUENCE FROM N.A.
RC Rounslev S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RC STRAIN=sv. Columbia;
RC STRAIN=sv. Columbia;
RA Lin X.;
RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RP (4)
RP SEQUENCE FROM N.A.
RC Rounslev S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RC STRAIN=sv. Columbia;
RC STRAIN=sv. Columbia;
RA Town C.D., Kaul S.;
RP Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RP (5)
RP SEQUENCE FROM N.A.
RC Rounslev S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RC STRAIN=sv. Columbia;
RC STRAIN=sv. Columbia;
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Ouedraogo C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RP (6)
RP SEQUENCE FROM N.A.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF367329; AK32916.1; -
DR EMBL; AC004111; AAC34227.2; -
DR EMBL; AY091698; AA010287.1; -
DR PIR; T02191; T02191.
DR GO; GO:0004497; Fmonooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR011128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PS00385; p450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 519 AA; 59603 MW; 38053C83F917457A CRC64;

Query Match 14.0%; Score 400.5; DB 10; Length 519;
Best Local Similarity 23.5%; Pred. No. 7.7e-21;
Matches 138; Conservative 110; Mismatches 231; Indels 109; Gaps 200

QY 1 MFVILVLTGALGIAPAFSWSIAFSLYLAPRRSLYNLQGNHNYFTGNF----- 51
DB 9 LVLVILF--IGIRFKAFMLVHPFVLRKLNQGISGPNY-RIFYGNLSEIKWKRE 64
QY 52 --LDLISARTGE-----EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHWKAYD 102
DB 65 SHLSILDPSSNDIFPRILPHYQKMSQYGETFLYWN-GTEPRICISDPELAKTLMLSNKL 123
QY 103 -YKPGWAAVRLIATGGVGTATSEGAHKRRIRIMTPSLSAQAQVSMVPIFLEKGMELVD 161
DB 124 PFVKSKARFENKVLGSLVFIQADWVRHRLINPAFSDRLKINTVWVDCTL----- 179
QY 162 KMMDEAAEKDMAVGESAGEKKATELETEGYVDVKDWVG----RATLDVWALAGFYKSDSL 217
DB 180 KMLEE-----WRKESTKEETHEPKIKKENNEBFQRLTIATISAFG----- 221

```

```

QY 218 QNKTNELYVAVGLTGDGAPATLDSFKALIMWDFVYFRTMKRRHEIPLTQGLAVSRVVGIE 277
Db 222 -----SSVVEGIEFVSQM-ELKRCYTTLSNQVSIPGTQYLPFPNSNIRVW 265
QY 278 LMEOKKQAVLGSASQAVDKDVGDRDLSLLVPEANLPSQKLSDEEVLQISNLL 337
Db 266 KLERKQNSIKRIISRLQSKSDYGDGLGILLKAYNTEG--KERRKSWIEIIEHCRTFF 323
QY 338 FAGYTSSTVLTMWPHRLSEDAVQDKLREI---CQIDTDMPTLDELNALPYLEAFVKE 394
Db 324 FGCHETTNNLAWTMTLLSLHQDWQEKLEBIFKECGKE-KTDSFESKJLKMNVIME 382
QY 395 SRLDPPSPYANRECLDKDEDFIPLAEFVIGRDGVS---INEVRITKGTWMLPLEFNRES 451
Db 383 SRLVGP-----VSALAREASVNTKIGDLIPKGTIVVPLLNKNSD 424
QY 452 KFIYEDAEFPERMLEDVDSLSNIEAPYGHQASFIISGPRACFGWRFVAEMKAFLEV 511
Db 425 KTLWGSADAKFNMFRAFNGVSRANHPNA---LLAFSVGPRACIGQNFVMEIAKTLTM 480
QY 512 TLRRVQFEFLISHPEYEH--ITLISRPVIGREKEGYQMRLOVKVE 557
Db 481 ILQRFRTSLCD--EYKTPVDNVTIQO-----YGLVMEQZPEE 518

RESULT 10
Q8LHV0 PRELIMINARY; PRT; 519 AA.
AC Q8LHV0;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN P0025D09.13.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Fiales; Poaceae;
OC Pharbitidae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cnv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone P0025D09."
RL Submitted (OCT-2001) to the ENBL/GenBank/DBJ databases.
CC 1. SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF004264; SACCIO362.1; -.
DR Gramene; Q8LHV0; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 519 AA; 58547 MW; 77E7B9BC1428FE38 CRC64;

Query Match 13.9%; Score 397.5; DB 10; Length 519;
Best Local Similarity 26.0%; Pred. No. 1.3e-20;
Matches 155; Conservative 80; Mismatches 225; Indels 137; Gaps 22;

QY 10 ALGLAAP-----SWASIAFFSLYLAP-----PRSSLYNLOGNHNVTFTGNFLDI 54
Db 2 ANGLLAWMVAAAAAVALASNAFSAVTLVWRFRAISRLRAQGVGGFGY-RFFSGNLGSI 60

QY 55 LSAR-----TGSEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHV 96
Db 61 KRFRGDGAGVNLVSSHDFLPVQPHFRKFIPLXGRTFLY-WFGAQNICLADVSVMVQV 119
QY 97 MKEAYD-IPKFGMAARVLRATGCVYTAEGEAHKRRRIMISLAQAQKSNVPIFEK 155

```

```

Db 120 LSDRTGIYPKNLTNPHFVRL-LGKGLVLTGDEWKRHRKVVPFAFNMDKLKMMT----- 172
QY 156 GMELVDKMMEDAAEKDMAVGESAGKATRLTEGVDVKDWGVGRATLDVMAAGF--DYK 213
Db 173 -----MTNSDCSRMSMSEWESELAAGGLVE---IELSRFEELTADVISHTAFGSSYK 223
QY 214 SDS---LQKNTNELYVAVGLT---DGFA---PTLDSFKALIMWDFVYFRTM-----KRRH 280
Db 224 EGKQVFLAQRELOFLAFSTFLTVQIPGFSYLPYTKNFKT--WSLDKKVRGMLMDIINTRH 281
QY 261 EPLATQGLAVSRVVGIELMEQKQAVLGSASDAQVDDKV--OGRDILSLVLRANTAAANL 318
Db 282 A-----NKDVAGYGNLGLMLSCAPEHG 306
QY 319 PSQKLSDEEVLQISNLLFAGYTSSTVLTMWPHRLSEDAVQDKLREIC-QIDTDM 377
Db 307 ESPQLSMDDEIIDCKTFFAGHDTTSHLLTWTMFLSTHPDWQEKLEBIEINMECQDKVP 366
QY 378 TLDELNALPYLEAFVKSRLRDPSPYANRECLDKDEDFIPLAEFVIGRDGVSINEVRITK 437
Db 367 TGDMLNKLKXVNMFLLETLRLYSPVSLIRKVDTDIE-----LGGIKMPE 411
QY 438 GTWVWMLPLENINRSKFIYGEDAEFRPERMLEDVDSLSNIEAPYGHQASFIISGPRACFG 497
Db 412 GALLTIPTATIHRRKEVWGEDADEFRPERFENGVTAAKHNA-----LUSFSSGPRSCIG 467
QY 498 WRFVAEMKAFLEVTLRVQF--EPIISHPEYEHITLISRPVIGREKEGYQMRLO 552
Db 468 QNFAMIEAKAVIAMILQRFSTLSPKVVHAPTDVITL-----RPKYGLPLMILK 515

RESULT 11
Q86SK3 PRELIMINARY; PRT; 503 AA.
AC Q86SK3;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Cytochrome P450.
GN CYP3A43/CYP3A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21839017; PubMed=11726664;
RA Finta C., Zaphiropoulos P.G.;
RT "Intergenic mRNA molecules resulting from trans-splicing."
RL J. Biol. Chem. 277:5882-5890(2002).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Zaphiropoulos P.G.;
RL Submitted (MAY-2003) to the ENBL/GenBank/DBJ databases.
DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 503 AA; 57394 MW; 4141F95B28B34A6E CRC64;

Query Match 13.8%; Score 394; DB 4; Length 503;
Best Local Similarity 28.7%; Pred. No. 2.2e-20;
Matches 142; Conservative 83; Mismatches 177; Indels 92; Gaps 20;

QY 47 FTGNFLDIISARTG-----TEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHV--KE 99

```

Db 46 FLGN---ILSHKGFMDMECHKKYGKVGPF---VDG---QQPVLAITDPDMIKTVLWKE 97
 QY 100 AYDY---PKGMAARVLRITAGDGVVTAEGEAHRRHRIIMPISLSAQAVKSMVPIFEKG 156
 Db 98 CYSVFTNRPRPGPVGFKSA---ISIAEDBEWKRLRLSPTFTSGKLKEMVPIIAQYG 153
 QY 157 MELVDKMMEDAAEKDVAESAGEKATRLTET-EGVDVKDVGWRATLDVMAAGADYKSD 215
 Db 154 DVLVRNLREA-----ETGKPTLKDVGAYSMOVIITSGVNIID 194
 QY 216 SLQKTNELVAVFGLTGDGAPILDSPKAIM-WDFV-PYFRWK-RRHEIPLTQGLAV-- 270
 Db 195 SLNNPOD-----PFVENTKLLRDFLDPPFLSITVFPFPIPILEVLNICY 240
 QY 271 -SRRVGIEMKQKQAVLGSASDAQVQKQVQGRDILSLVRANIAANLPSCKLSDEV 329
 Db 241 FPREVNTFLRSVKR-----MKESRLBDTQKRVDFLQLMDSQSKETESHKALSDEL 295
 QY 330 LAQISNLLFAGYETSSVLTWFMHRLSEDKAVODKLEESICQI--DTMPTLDELNALPY 387
 Db 296 VAQSTIFIFAGYETSSVLTWFMHRLSEDKAVODKLEESICQI--DTMPTLDELNALPY 387
 QY 388 LEAFVKESLRDPPSPVANRECKLDEDFIPLAEFVIGRDSVINEVRITKGMVMLPLFN 447
 Db 356 LDMVNETLRLFPAMRLERVKCKDVE-----INGMFIKGVVWIPSYA 400
 QY 448 INRSKFIYGEAEPRFRERMLVEDVTDLSNS-IEAPYGHQASFGSPRACFCGRFAVAMK 506
 Db 401 LHRDP-KWYTEPEKFLPERFSKXKNDIDPYIYTPFG-----SGPRNCIGMRFFALNMK 453
 QY 507 AFLFVTLRRVQFEP 520
 Db 454 LALILVQLNFSFKP 467

RESULT 12

Q9FEEL PRELIMINARY; PRT; 544 AA.
 AC Q9FEEL1;
 DT 01-MAR-2001 (TremBLrel. 16, Created)
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 GN P068A04.9 OR P0006C01.24.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P068A04.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0006C01.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -/- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF002839; BAB19103.1; -
 DR EMBL; AF002744; BAB19082.1; -
 DR Gramene; Q9FEEL1;
 DR GO; GO:000497; P:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; p450.1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.

KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 544 AA; 60867 MW; 273EAF5968D1A024 CRC64;
 Query Match 13.8%; Score 393.5; DB 10; Length 544;
 Best Local Similarity 26.4%; Pred. No. 2.7e-20;
 Matches 159; Conservative 83; Mismatches 215; Indels 145; Gaps 25;
 QY 6 LLTGALGLAFAFSAWIAFFSLYLAPRR-SSLYNLQGNHTNY-----FTGNFLDILGAR 58
 Db 31 LLGNAVALLLVNAQMLEWALAPRMEBALAQGLRGTYQYRFLHGLDTEDLRLVTAAR 90
 QY 59 TG-----EBHAKYREKYGSTLRFPAGIAGAPVLNSTDPKVFHV 96
 Db 91 SKPQFMDRPHDFIPRVAPLLHRALEHGR-----VSFTWFGPMPTVITDPLVREV 142
 QY 97 MKEAV-DYPKQMAARVLRITAGDGVVTAEGEAHRRHRIIMPISLSAQAVKSMVPIFEK 155
 Db 143 LSNKFGHEPKTKLATRLSKLLVG-GLVILHGEKWKVHRIMNPAFHAELKELKMLPAFSAS 201
 QY 156 GMLVDKMMEDAAEKDVAESAGEKATRLTETEGVDVKDVGWRATL--DYMALAGFDYK 213
 Db 202 CSELIGRW-----ENAVAASVQKAE-----LDI--WPDFQNLSGDVISRRAAFGVR 244
 QY 214 SDS-----LQNKTNELVAVFGLTGDGAPILDSPKAIMMDFVPVF-----RTMKRRHEI 262
 Db 245 HHEGRQIFLLQAEQARLV-----QSFS--NVIPLSYENNRMK----- 283
 QY 263 PLTQGLAVSRV-----GILMEQKQAVLGSASDAQVQKQVQGRDILSLVRANIAAN 317
 Db 284 -----AIDREIKSILRGI-IEKQKATKNGEAS-----KD-----DLGLLLQGNMDY 326
 QY 318 LPESOKLSD---BEVLAQISNLLFAGYETSSVLTWFMHRLSEDKAVODKLEESICQI- 372
 Db 327 SDEGSKSGKMTVEIIDECKLFYFAGMETTAVLLTWTWVALSMHPEDQARREBILQVF 386
 QY 373 DTDMPITDELNALPYLEAFVKESLRDPPSPVANRECKLDEDFIPLAEFVIGRDSVINE 432
 Db 387 GRNKEDINGVSRKLVTVTLHEVLRVPPVVMNRRTYKEIE-----LGG 431
 QY 433 VRITKGTVMPLFNINRSKFIYGEAEPRFRERMLVEDVTDLSNSIEAPYGHQASFGSP 492
 Db 432 VRYPAQVMSLPVLFIRHDAAGHGDAGEFDPGRFAEGVARACKDPGA--GAFFPFSWGP 489
 QY 493 RACFCWRFAVAMKAFLEVTILRRVQPE--PLIHPHYEHITLIISRPVGRKEGYQWR 550
 Db 490 RICIQONFALLEAKVAGLMILQRFAPFELSPAYAHAPYTVITL-----HPQHGVPVR 540
 QY 551 LQ 552
 Db 541 LR 542

RESULT 13

Q8CJF2 PRELIMINARY; PRT; 497 AA.
 AC Q8CJF2;
 DT 01-MAR-2003 (TremBLrel. 23, Created)
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE Cytochrome P450 3A.
 GN CYP3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Matsubara T., Nagata K., Yamazoe Y.;
 RT "Isolation and characterization of a novel rat CYP3A form.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB084894; BAC23085.1; -
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.

[illegible]

Search completed: April 2, 2004, 14:00:40
Job time : 41.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 11 Seconds
(without alignments)
2636.643 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLLTGALGLAFAFWAS.....RIVGREXEGYQMLQKVPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414	14.5	503	1 CP3T_PIG	P79401 sus scrofa
2	413	14.5	503	1 CP39_RAT	P1538 rattus norv
3	408.5	14.3	503	1 CP30_SHEEP	Q29496 ovis aries
4	404	14.2	503	1 CP3C_CANFA	P24463 canis famil
5	402	14.1	501	1 CP36_RABIT	P11707 oryctolagus
6	399	14.0	503	1 CP33_HUMAN	P05184 homo sapien
7	398	14.0	503	1 CP3D_MOUSE	Q84484 mus musculu
8	396.5	13.9	502	1 CP34_HUMAN	P08684 homo sapien
9	395.5	13.9	503	1 CP38_WACEA	P33268 macaca fasc
10	394	13.8	504	1 CP3B_MOUSE	Q54459 mus musculu
11	391	13.7	504	1 C341_MOUSE	Q9jma7 mus musculu
12	385.5	13.5	503	1 CP37_HUMAN	P24462 homo sapien
13	385	13.5	502	1 CP35_HUMAN	P20815 homo sapien
14	385	13.5	507	1 CP3S_BOVIN	P79102 bos taurus
15	377.5	13.2	503	1 CP3L_CALJA	O18993 callitrix
16	371.5	13.0	501	1 CP3V_WESAU	O70537 mesocricetu
17	370	13.0	504	1 CP31_RAT	P04800 rattus norv
18	368	12.9	504	1 CP3G_MOUSE	Q64481 mus musculu
19	362	12.7	503	1 CP3F_CAVPO	Q64406 cavia porce
20	359.5	12.6	511	1 CP30_FUNHE	Q64462 mus musculu
21	356.5	12.5	496	1 CP33_FUNHE	Q9pve8 fundulus he
22	355.5	12.5	503	1 CP3E_CAVPO	Q64417 cavia porce
23	351.5	12.3	496	1 C356_FUNHE	Q8axys fundulus he
24	351	12.3	520	1 CP38_HUMAN	P38187 homo sapien
25	350.5	12.3	1054	1 CYPE_BACSU	O08336 bacillus su
26	350	12.3	504	1 CP32_RAT	P05183 rattus norv
27	350	12.3	526	1 CP35_RAT	P51870 rattus norv
28	349.5	12.3	524	1 CP31_RAT	P33274 rattus norv
29	348.5	12.2	511	1 CP4B_HUMAN	P13584 homo sapien
30	346.5	12.1	511	1 CP4B_RAT	P15129 rattus norv
31	346.5	12.1	524	1 CP3C_HUMAN	Q2hcs2 homo sapien
32	345.5	12.1	520	1 YRV2_CAEEL	Q27514 caenorhabdi
33	344.5	12.1	520	1 C340_ORILA	Q98t91 oryzias lat

34	344	12.1	520	1 CP3B_HUMAN	Q08477 homo sapien
35	344	12.1	524	1 CP3B_HUMAN	Q9hb16 homo sapien
36	340.5	11.9	524	1 CP72_CATRO	Q05047 catharanthu
37	340	11.9	503	1 CP3P_MOUSE	Q09158 mus musculu
38	338.5	11.9	503	1 C343_HUMAN	Q9hbs5 homo sapien
39	337	11.8	496	1 C4AE_DROME	Q46054 drosophila
40	337	11.8	503	1 CP3H_CAVPO	Q64409 cavia porce
41	337	11.8	520	1 CP32_HUMAN	P78329 homo sapien
42	336.5	11.8	518	1 CP3R_ONCMY	O42563 oncothyachu
43	336	11.8	522	1 CP34_RAT	P51869 rattus norv
44	334.5	11.7	509	1 YRV4_CAEEL	Q27516 caenorhabdi
45	334.5	11.7	580	1 C973_ARATH	O23365 arabidopsis

ALIGNMENTS

RESULT 1

CP3T_PIG
ID_CP3T_PIG STANDARD; PRT; 503 AA.
AC P79401, 1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DI 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A29 (EC 1.14.14.1) (CYP11A29).
GN CYP3A29.
OS Sus scrofa (Pig).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=98347363; PubMed=9682441;
RA Nissen P.H., Winteroe A.K., Fredholm M.;
RT "Mapping of porcine genes belonging to two different cytochrome-P450
subfamilies.";
RL Anim. Genet. 29:7-11(1998).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
tissues by various foreign compounds, including drugs, pesticides,
and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z93099; CAB07513.1; -.
HSSP; P14779; IJPTZ.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
KW METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57198 MW; E3D411B2674FD17F CRC64;

Query Match 14.5%; Score 414; DB 1; Length 503;

Best Local Similarity 27.4%; Pred. No. 8.9e-17;
Matches 158; Conservative 85; Mismatches 197; Indels 136; Gaps 23;

QY 12 GLAFAWSAIA--PESLYLAPRRS-----SLYNLQGNHNTWYFTGNFLDLSARTGEH-- 63
Db 6 GSTETWLLATSLVLLLYTYSHGLFKLGLTGPRLPYF-GN---ILYRKGVDPD 61
QY 64 AKYREKYGSTURFAGI--AGAPVUNSTDPKVFHVM-KEAY-----DYPKPGMAARVL 113
Db 62 KCFQOYQ---KMGVYDGRQFLAVTDPMNKSVLVKECYSVFTNRRRFGGLGAMRNAL 118
QY 114 RATGDDGVVTAAGGAHKKHRRIMPSLSAQAQVSWVPFLEKGMELVDQWEDAAEKDWA 173
Db 119 SL-----AEDBEWKEIRLLSPTFTSGKLKEMFPLISHYGDLLVSNLKE-AEKG-- 167
QY 174 VGESAGEKATRLTEGVVDVQVGRATLDYMALAGFDYKSDSLQNK-----TNELYVAF 228
Db 168 -----KPVTKMDIFGAYSNDVITSTAFGVNIDSLNNPQDPFFVENSKKLLK 212
QY 229 VGLTDFGAPTLDLDSFKAIMWDF-----VPYF-RTMKRHEIPLTQGLAVSRVVG 275
Db 213 FSFFDPFLLSLFPPLTPIEVNLITLFPKSSWNFTKSVKEMKESRLT----- 262
QY 276 IELMQKQAVLGASDAQVQKQVQGRDILSLVLRANTAAANLPESQKLSDEVLQAQSN 335
Db 263 -----DQCKRRV-----DLLQNLNSQNSKENDPHKLSNBEELVAQGI 301
QY 336 LUFAGYETSSTVLTWFMHRLSEDAVQDKLREICQI--DTDMPTLDELNALPYLEAFVK 393
Db 302 FIFAGYETTSALSLLAVELATHPDVQKQJEEIATEFPKAPTVDALQAQVEYLDVWN 361
QY 394 ESRLDPSPPVNRCLKDEDFIPLAEFVIGRQSVINEVRITKGTWMLPFLFNINRKF 453
Db 362 EYRLPYTAARLEACKQDVE-----IHGVFVPGKVVVVVVFVLRDPD 406
QY 454 IYGEDAEFRPERMLEDVTDLSNIEAPYGHQASFSISGRACPMRFAVAENKAFIFVTL 513
Db 407 LWPE-PEEFPERESKHKDTIN---PYTY-LPFGTGPENCIGMRFAALMNMKLAVRL 460
QY 514 RVQVEPI-----ISHPEYEHITLISR 536
Db 461 QNFSFKPKCTQIPUKLTQGLTQPERKPVVILKILR 496

RESULT 2

CP39 RAT STANDARD; PRT; 503 AA.
AC P51578; O64557; O64631;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome P450 3A9 (EC 1.14.14.1) (CYFIIA9) (P450-OLF3) (Olfactive) (3AH15).
DE GN CYP3A9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97144501; PubMed=8990268;
RA Mahne A., Strotkamp D., Roos P.H., Hanstein W.G., Chabot G.G.,
RA Nef P.;
RT "Expression and inducibility of cytochrome P450 3A9 (CYP3A9) and
RT other members of the CYP3A subfamily in rat liver.";
RL Arch. Biochem. Biophys. 337:62-68 (1997).
RL [2]
RP SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=96220175; PubMed=8660328;
RX Wang H., Kawashima H., Strobel H.W.;
RA "CDNA cloning of a novel CYP3A from rat brain.";

RL Biochem. Biophys. Res. Commun. 221:157-162 (1996).
CC -!- FUNCTION: This isozyme seems to be implicated in oifaction. Active
CC in the demethylation of erythromycin as well as benzphetamine.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: Mainly expressed in olfactory epithelium.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U60085; AAB03662.1; -;
DR EMBL; U46118; AAC52582.1; -;
DR PIR; JC4702; JC4702.
DR HSP; P14779; 1JPZ.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00667; P450; 1; P4501CYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR OXidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Olfaction.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 457 457 F -> V (IN REF. 2).
SQ SEQUENCE 503 AA; 57811 MW; 9BBB13E690675EB4 CRC64;

Query Match 14.5%; Score 413; DB 1; Length 503;
Best Local Similarity 27.4%; Pred. No. 1e-16;
Matches 145; Conservative 86; Mismatches 178; Indels 120; Gaps 22;

QY 54 ILSARTG-EHAKY-REKYGSTLRFAGI--AGAPVUNSTDPKVFHVM-KEAY----- 101
Db 50 ILAYKGFWEFDKCHKYK--KLWGLYDGRQVLAITDPIIKTVLKECYSTFNR 106
QY 102 DYPKGMARVLRATGDGVVTAAGBAHKKHRRIMPSLSAQAQVSWVPFLEKGMELVD 161
Db 107 NFGPVGILKXAI-----SEDEWQIRALLSFTTSGKLKEMFII----- 149
QY 162 KQMEDAAEKQNAVGESAGEKATRLTEGVVDVQVGRATLDYMALAGFDYKSDSLQNK 221
Db 150 NQYTDMLVRNMQSGSEKGP-----TSMKDIAGAYSDVITATSGVNVDSLNNPQ 200
QY 222 NELYVAFVGLTGDGAPTLDLDSFKAIM-WD-FVPYFRTMKRHEI-PLTQGLAYSR----- 272
Db 201 D-----PFVEKVKLLKFDIFDPLFLSVTLFPFPLTFLFEALNVSMFPRDI 246
QY 273 ---RVGIELMEQKQAVLGASDAQVQKQVQGRDILSLVLRANTAAANLPESQKLSDEEV 329
Db 247 DFFKTSVERMKENR-----MKEKEQRQVDFLQMLNSQNSKVXKDSHKALSDVEI 295
QY 330 LAQISNLFPAGYETSSTVLTWFMHRLSEDAVQDKLREICQIDTMP-----TLDELNA 384
Db 296 VAOQVIFIFAGYETTSALSFLVLLAHDPDIQKQJDE---IDAAALPNKATYDTLLQ 352
QY 385 LPYLEAFVKESLRLDPPSPVNRCLKDEDFIPLAEFVIGRQSVINEVRITKGTWMLP 444
Db 353 MEYLDWVWNETLRLVPIAGRLERLVCKTDE-----INGVFIPKGTVMIP 397
QY 445 LFNINRSFITYGDAEPRPERMLEDVTDLSNIEAPYGHQASFSISGRACPMRFAVAE 504
Db 398 TFLHKDHPH-YMPSEPERPERFSKQNDNIN---PYMY-LPFGNGPNCIGMRFAALM 451
QY 505 MKAFLEVTILRRVQVEPI-----ISHPEYEHITLISRPIV 540
Db 452 MKVALFRVLQNPSFOPCKETQIPKLKQGLQPEKFLLLKVVSRDET 500

```

RESULT 3
ID CP30 SHEEP STANDARD; PRT; 503 AA.
AC Q29456;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A24 (EC 1.14.14.1) (CYP3A24).
GN CYP3A24.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Ching M.S., Chun-Jing J., Ghabrial H., Wooley P.J., Smallwood R.A.,
RA Morgan D.J.;
RL "Ovine foetal liver CYP3A24."
RT Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U59378; AB02657.1; -
CC HSP; P14775; IJPE2.
CC InterPro; IPR001128; Cytochrome_P450.
CC DR Pfam; PF0067; P450; 1.
CC DR PRINTS; PR01689; EP4501ICYP3A.
CC DR PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC Microsome; Endoplasmic reticulum.
CC FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC SQ SEQUENCE 503 AA; 57360 MW; 804223EAD0304238 CRC64;

Query Match 14.3%; Score 408.5; DB 1; Length 503;
Best Local Similarity 28.5%; Pred. No. 1.8e-16;
Matches 165; Conservative 82; Mismatches 197; Indels 135; Gaps 27;

QY 7 LTGALGLAFAFNASIAFFSLYLAPRRS-----SLYNLQGNHTNYFTGNFLDILSARTG-- 60
DB 3 LTPSFLTWLVLLAISLVLLLYGYSHGLFKLVGSGRPLPYF-GN--VLSRYKGVG 58
QY 61 ----BEHAKYREKYGSTLRPAGIAGAPVNLSTDPKYNHVM-KEAYD-----YKPGM 108
DB 59 EFDEECFKYKQWGVG---FEG--KQPLLVITDPDVIKTVLVKCYSVFTNRRVFGPMGI 113
QY 109 AARVLRIATGDGVTAEGEAKHRRIRIMPSLSQAVKSMVPIFEKGMELVDKMMEDAA 168
DB 114 MK-----NAVVAEDQWKRIRITLSPFTSGKLDMPFLIGKYGDVLRNLRKE-A 164
QY 169 EKDMAVGESAGEKKATRLTEGTGVVDVQWVGRATLDWMALAGFDYKSDLSQNKINELYAF 228

```

```

DB 165 EK-----GKS-----VNMKIDFGAYSMVDVITSTSGVNDISLGNP----- 199
QY 229 VGLTDGFAPT-----LDSP--KAIMWDF--VPYFR-----TMKRRHEIP--LTQGLAVS 271
DB 200 ---QDPVENAKLLRNFNLDLPFLSVLFPFLPIFEVLNITMFPKSAVDFTK--SYK 254
QY 272 RRVGIELMECKQAVLGSASDQAVDKQVQGRDILSLVRANIAANLPESQKLSDEEVLA 331
DB 255 RIKESRLKQKQKPRV-----DFQLMINSONSKETDNHKKALSQOELMA 297
QY 332 QISNLLPAGYETSTVLTMWHLRSLSDKAVQKLRKEICQIDTDM-----PTLDELNALIP 386
DB 298 QSVIFIFAGYETTSNTLSFLLYILATHPDVQKQLOBE---IDATFPNKAPPTDYDLAQME 354
QY 387 YLEAFVKESRLDPPSPFYANRECKLDEDFIPLAEPVIGRSGVINEVRITKGTVMVLP 446
DB 355 YLDMVWNETLUMFPIAVLRDLRCKKVE-----IHGVSPKGTAVTVPIF 399
QY 447 NINRSKFIYGEDAEFRPERLEWEDVTDLSNS-IEAPYGHQASFIQGFACGFRFAVAM 505
DB 400 VLHRDQPLWPE-PEERPERPFSKKNKDSINPYVLPFG-----TGNRCNIGMRFAIMNM 452
QY 506 KAFLEVTILRRVQPEPIISHPEYEHITLISRPPIVGRK 544
DB 453 KLAIVRVQLQNFSPKPC---KETQIFPKINSQGLIRPEK 497

RESULT 4
CP3C CANFA
ID CP3C CANFA STANDARD; PRT; 503 AA.
AC P24463;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A12 (EC 1.14.14.1) (CYP3A12).
GN CYP3A12.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RX MEDLINE=91159488; Pubmed=2001406;
RA Ciaccio P.J., Graves P.E., Bourque D.P., Glimmann-Gibson B.,
RA Halpert J.R.;
RT "cDNA and deduced amino acid sequences of a dog liver cytochrome
RT P-450 of the IIIa gene subfamily."
RL Biochim. Biophys. Acta 1088:319-322(1991).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54915; CAA38687.1; -
CC PIR; S14275; S14275.

```


Db 322 HPDVQKQKQEE---IDTLLENKELATYDILVRMEYLDVMVNETLRLPIAGRLERVKCKD 378
Qy 413 EDPIPLAEPVIGRDGVSINVRITKGTVMVPLNFNINRSKFIYGEDAEFRPERWLEDVT 472
Db 379 VD-----INGTFIPKGTIVMPTALHRRDPQWTE-PDBERBERSKNK 422
Qy 473 DSLNS-TEAPYGHQASFISSPRACFGWRFAVEMKAFVLTLLRVQFE-----PI--- 521
Db 423 DNINPYIYHFFG-----AGPRCLGMRFAIMNKLALVRLMNFSFKLCKETQVPLKLG 476
Qy 522 ---ISHPEYEHITLIISRPRIV 540
Db 477 KQGLQPEKEIVLKVWSDGII 498
Db 522 ---ISHPEYEHITLIISRPRIV 540
Qy 477 KQGLQPEKEIVLKVWSDGII 498

RESULT 6
CP33 HUMAN
ID CP33 HUMAN STANDARD; PRT; 503 AA.
AC P05184;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A3 (EC 1.14.14.1) (CYP11A3) (Hlp).
GN CYP3A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=86259780; PubMed=3460094;
RA Molowa D.T., Schuetz E.G., Wrighton S.A., Watkins P.B., Kremers P.,
RA Mendez-Picon G., Guzelian P.S.;
RT "Complete cDNA sequence of a cytochrome P-450 inducible by
RT glucocorticoids in human liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5311-5315 (1986).
RN [2]
RP SEQUENCE OF 1-20.
RC TISSUE=Liver;
RX MEDLINE=86298342; PubMed=3898085;
RA Watkins P.B., Wrighton S.A., Maurel P., Schuetz E.G., Mendez-Picon G.,
RA Parker G.A., Guzelian P.S.;
RT "Identification of an inducible form of cytochrome P-450 in human
RT liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6310-6314 (1985).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By glucocorticoids.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: D00003; BAA0001.1; -
DR EMBL: M13785; AAA35742.1; -
DR PIR: A29410; A29410.
DR HSSP: P14779; LJPZ.
DR Genew; HGNC:2636; CYP3A3.
DR GO; GO:0005792; C:mitochondrion; TAS.
DR GO; GO:0015034; F:cytochrome P450 activity; TAS.
DR InterPro; IPR001128; Cytochrome_P450.

DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00667; P450; 1.
DR PRINTS; PRO1689; EP450IICYP3A.
DR PRINTS; PRO385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT INIT MET 0
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57428 MW; 9885D1F729658FPC0 CRC64;
Query Match 14.0%; Score 399; DB 1; Length 503;
Best Local Similarity 28.4%; Pred. No. 6.4e-16;
Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;
Qy 13 LAAPSW--ASIAFFSYLAPRES-SLYNLQG--PNTNY--FTGNFLDILSARTG-----E 61
Db 6 LAMETWLLAVSLVLLYLYGTHSHGUFKGLGPGTLPFLGN-----ILSYHKGFCMPDME 62
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVENHYM--KEAYD-----YKPGVAAARVL 113
Db 63 CHKYGKVGWF---YDG--QQPVLAITDPDMIKVLVKECYSVFTNREPPFGVGFMSAI 117
Qy 114 RIATGCVTAEGEAKHRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDWA 173
Db 118 SI-----AEDEWKRLRLSPTFTTSGLKWKVPPIAQYGDVLRNLRRE----- 162
Qy 174 VGESAGEKKATRLTEGVVDKDWGRATLDVNMALAGFDYKSDSLQNKTNELYVAFVGLTD 233
Db 163 -----RETGKPVTLKDVFGAYSMDVITSSSGVNVDSLNNPD----- 200
Qy 234 GPAPTLDSPKAIM-WDFV--PYFRTMK--RRHEPLPLQGL-----AVSRVG 275
Db 201 ---PLVENTKKLRFDFLDFFLSITVFPFLPILEVLNLCVFPEVNTNFKAVKRWKE 257
Qy 276 IEAMECKQAVLGASDAQVDKQVGRDILSLVRANIAANLPESQK--LSDEEVLAQIS 334
Db 258 SRLDETQKRV-----DFLQIMDSHKNSKETESHKALSLELVAQSI 300
Qy 335 NLLFAGYETSTVLTWMFHLRSLSDKAVQDKLREICQI--DTDMPFLDELNALPYLEAFV 392
Db 301 IFIFAGYETTSVLSFIMVELATHPDVQKLOBEIDAVLPNKAPPTDYDTVLMQEYLDMVV 360
Qy 393 KESLRDPPSPYANRECLKDEDFIPLAEVPIGRDGSVINEVRIKGTVMVPLNFNINRSK 452
Db 361 NETLRLPFIAMRLERVCKDVE-----INGMFIKGVWVMIPISYALHRDP 405
Qy 453 FIYGEDAEERPERWLEDVTDSLNS-IEAPYGHQASFISSPRACFGWRFAVEMKAFV 511
Db 406 -KYWTEPEKFLPERFSKKNKDIDPIYTFPG-----SGPRNCIGMRPALNMKLALIR 458
Qy 512 TLRRVQFEP 520
Db 459 VLQWFSEKP 467
RESULT 7
CP3D_MOUSE
ID CP3D_MOUSE STANDARD; PRT; 503 AA.
AC Q64464;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A13 (EC 1.14.14.1) (CYP11A13).
GN CYP3A13 OR CYP3A-13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ddy; TISSUE=Liver;
RX MEDLINE=95101705; PubMed=7803471;

RA Siest J.-P., Huang J.-D.;
RT "Novel mutations of CYP3A4 in Chinese."
RL Drug Metab. Dispos. 29:268-273 (2001).
RN [8]
RC TISSUE=Liver;
RP SEQUENCE FROM N.A.
RA Zhuge J., Qian Y., Xie H., Yu Y.;
RT "Sequence of a new human cytochrome P450-3A4 cDNA."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP CHARACTERIZATION.
RA MEDLINE=97243737; PubMed=9088578;
RX Zhang H., Coville P.F., Walker R.J., Miners J.O., Birkett D.J.,
RT Wanwimolruk S.;
RA "Evidence for involvement of human CYP3A in the 3-hydroxylation of
RT quinine."
RL Br. J. Clin. Pharmacol. 43:245-252 (1997).
RN [10]
RP CHARACTERIZATION.
RA MEDLINE=98122830; PubMed=9456308;
RX Zhao X.J., Kawashiro T., Ishizaki T.;
RT "Mutual inhibition between quinine and etoposide by human liver
RT microsomes. Evidence for cytochrome P4503A4 involvement in their
RT major metabolic pathways."
RL Drug Metab. Dispos. 26:188-191 (1998).
RN [11]
RP VARIANTS PRO-221 AND THR-444.
RA MEDLINE=20132287; PubMed=10668853;
RX Sata F., Sapone A., Elizondo G., Stocker P., Miller V.P., Zheng W.,
RA Raunio H., Crespi C.L., Gonzalez F.J.;
RT "CYP3A4 allelic variants with amino acid substitutions in exons 7 and
RT 12: evidence for an allelic variant with altered catalytic activity."
RL Clin. Pharmacol. Ther. 67:48-56 (2000).
RN [12]
RP VARIANTS GLN-161; SER-188; PRO-292; THR-444 AND SER-466, AND
RP CHARACTERIZATION OF THE VARIANTS
RX MEDLINE=21571796; PubMed=11714865;
RA Dai D., Tang J., Rose R., Hodgson E., Bienstock R.J.,
RA Mohrenweiser H.W., Goldstein J.A.;
RT "Identification of variants of CYP3A4 and characterization of their
RT abilities to metabolize testosterone and chlorpyrifos."
RL J. Pharmacol. Exp. Ther. 299:825-831 (2001).
RN [13]
RP VARIANTS ASP-55; GLN-129; ILE-169; HIS-173; MET-362; PHE-372 AND
RP LEU-415.
RX MEDLINE=21364010; PubMed=11470997;
RA Eiselt R., Domanski T.L., Zibat A., Mueller R., Presecan-Siedel E.,
RA Huestert E., Zanger U.M., Brockmoller J., Klenk H.-P., Meyer U.A.,
RA Khan K.K., He Y.-A., Halpert J.R., Wojnowski L.;
RT "Identification and functional characterization of eight CYP3A4
RT protein variants."
RL Pharmacogenetics 11:447-458 (2001).
RN [14]
RP VARIANTS PRO-14; GLN-161; HIS-173; SER-184 AND PHE-372.
RX MEDLINE=21864884; PubMed=11875366;
RA Lamba J.K., Lin Y.S., Thummel K., Daly A., Watkins P.B., Strom S.,
RA Zhang J., Schuetz E.G.;
RT "Common allelic variants of cytochrome P4503A4 and their prevalence in
RT different populations."
RL Pharmacogenetics 12:121-132 (2002).
RN [15]
RP FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It performs a variety
CC of oxidation reactions (e.g. caffeine 8-oxidation, omeprazole
CC sulfoxidation, midazolam 1'-hydroxylation and midazolam 4-
CC hydroxylation) of structurally unrelated compounds, including
CC steroids, fatty acids, and xenobiotics. The enzyme also
CC hydroxylates etoposide.
CC
CC -!- CATALYTIC ACTIVITY: Quinine + NADPH + O(2) = 3-hydroxyquinine +
CC NADP(+) + H(2)O
CC
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC
CC -!- TISSUE SPECIFICITY: Expressed in prostate and liver.
CC
CC -!- INDUCTION: P450 can be induced to high levels in liver and other

tissues by various foreign compounds, including drugs, pesticides,
and carcinogens.
CC
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP3A4 alleles;
CC WWW="http://www.imm.ki.se/cypalleles/cyp3a4.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; M18907; AAA35745.1; --
CC EMBL; M14096; AAA35744.1; --
CC EMBL; X12387; AAA30944.1; --
CC EMBL; J04449; AAA35747.1; --
CC EMBL; AF182273; AAF13598.1; --
CC EMBL; AF280107; AAG32290.1; --
CC EMBL; AF209389; AAP21034.1; --
CC PIR; A29815; A29815.
CC HSPF; P14779; IUPZ.
CC
CC
CC MIN; 124010; --
CC GO; GO:0015034; F:cytochrome P450 activity; TAS.
CC GO; GO:0004497; F:monooxygenase activity; TAS.
CC GO; GO:0006629; P:lipid metabolism; TAS.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008072; EP450_CYP3A.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PRO1689; EP450IICYP3A.
CC PRINTS; PRO0385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; NADP; Endoplasmic reticulum; Polymorphism.
FT INIT_MET 0 0 BY SIMILARITY.
FT METAL 441 441 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT VARIANT 14 14 L -> P (in allele CYP3A4*14).
FT VARIANT 55 55 G -> D (in allele CYP3A4*7).
FT VARIANT 117 117 I -> V (in allele CYP3A4*4).
FT VARIANT 129 129 R -> Q (in allele CYP3A4*8).
FT VARIANT 161 161 R -> Q (in allele CYP3A4*15).
FT VARIANT 169 169 V -> I (in allele CYP3A4*9).
FT VARIANT 173 173 D -> H (in allele CYP3A4*10).
FT VARIANT 184 184 T -> S (in allele CYP3A4*16).
FT VARIANT 188 188 F -> S (in allele CYP3A4*17; exhibits
FT lower turnover numbers for testosterone
FT and chlorpyrifos).
FT FTID=VAR_014322.
FT P -> R (in allele CYP3A4*5).
FT FTID=VAR_011605.
FT
Query Match 13.9%; Score 396.5; DB 1; Length 502;
Best Local Similarity 28.7%; Pred. No. 8.9e-16;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
QY 13 LAAPSW-ASIAFFSLYLAPRRS-SLYNQO-PNHTNY-FTCNFLDILSARTG-----E 61
Db 6 LAMETMLLAVALSLVLLYLXGTHSHGLFKKLGKLPGETPLPFLGN---ILSYHKGFQCFQME 62
QY 62 EHAKYREKVGSTLRAGIAGAPVLNSTDFKVFHNM-KEADYD---PKPGMAARVLRIAT 117
Db 63 CHKIKYKVGWF---YDG--QQPVLAITDPMKTVLKVKECYSVFTNRRPFGPGFMKSA- 116

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-eib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR ENBL, S53047; AAS24952.1; -

DR PIR; S28168; S28168.

DR HSP; P14779; LJP2.

DR InterPro; IPR001128; Cytochrome P450.

DR InterPro; IPR008072; EP450_CYP3A.

DR Pfam; PF00067; P450; 1.

DR PRINTS; PR01689; EP450LICYB3A.

DR PRINTS; PR00385; P450.

DR PROSITE; PS00066; CYTOCHROME_P450; 1.

DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;

KW Microsome; Endoplasmic reticulum.

KW METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

SQ SEQUENCE 503 AA; 57511 MW; D701B6FB83AC8FB CRC64;

Query Match 13.9%; Score 395.5; DB 1; Length 503;

Best Local Similarity 28.6%; Pred. No. 1e-15;

Matches 155; Conservative 85; Mismatches 187; Indels 115; Gaps 26;

QY 13 LAAFSW--ASTAFSLYLAPRR--SLYNLOG-PNHNY-FTGNFLDLSARTG-----E 61

DB 7 LAVETWLLAVTLVLLYLYGTHSHGLFKLGIPGPTPLPLGN---ILSYRKGFWTFDME 63

QY 62 EHAKYREKYGSTLAPAGIAGAPVLNSTDPKVFHVM--KEAYD-----YKPGMAARVL 113

DB 64 CYKKGKVGWGP---YDG--RQPLAITDPNIXITLVKCEYSVFTNRRPFGVGMKNAI 118

QY 114 RATGSGVVTAEGEAHKKRRIMIPSLAQAQVSKMVPFLEKGMELVDKQMEDAEDAKDMA 173

DB 119 SI-----AEDBEWKIRSLSPFTSGKLEKEMVPIIAKYGDVLRNLREA----- 164

QY 174 VGESAGEKKATRLT-EGVDVYKDWVGRATLDVMAAGFDYKSDSLQNTNELYAVFVGLT 232

DB 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVWIDSLNPD----- 201

QY 233 DGFAPTLDSFALM-WDFV-PYFRTMK-RRHEPLTQGLAVS---RRVGLMECKQKQAV 286

DB 202 -----PFVENTKLLRFDFLDPFLSITFFPIILEVLNISI:FPREVTSFLRKSVKRI- 256

QY 287 LGSASDAQVKKDVQGR--DITSLLVRAANTAAALPESQKLSDEEVLQATSNLLFAGYETS 344

DB 257 -----KESRLKDTQKRVDFLQMLDWSQNSKETESHKALSDLELVAQSIIFAGYEIT 310

QY 345 STVLTWMEFRLSEDKAVQDKLREICQIDTDM-----PTLDELNALPYLEAFVKESLRLD 399

DB 311 SSVLSFIYELATHPDVQQLQEE---IDTVLPNKAPPTVTVLQMEYLDVMVNETLRF 367

QY 400 PPSPYANRECLKDEDFIPLAEFVIGRDSGVINEVRITKGTMTVMLPLFNINRSKFIYGEDA 459

DB 368 PIAMRLERVCKKQVE-----INGIFIPKGVVWVMPYSALHHPD-KYWPPEP 411

QY 460 EEFREPERLEEDVTDLSNS-TEAPYHQASISCPACFGWRFAVEMKAFLEVTLLRRVQF 518

DB 412 EKFLERFSKKNNDNIDPIYITPFG-----SGPRNCIGRFAIMMKLAIIRVLQNFSP 465

QY 519 EP 520

DB 466 KP 467

RESULT 10

ID CP3B_MOUSE STANDARD; PRT; 504 AA.

AC 064459;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cytochrome P450 3A11 (EC 1.14.14.1) (CYP11A11) (P-450IIIA11) (P-450UT)

GN CYP3A11 OR CYP3A-11.

QY 118 GGVVTAEGEAHKKRRIMIPSLAQAQVSKMVPFLEKGMELVDKQMEDAEDAKDMAVGES 177

DB 117 ---ISIAEDWKLRLSLSPFTSGKLEKEMVPIIAKYGDVLRNLREA----- 163

QY 178 AGEKATRLT-EGVDVYKDWVGRATLDVMAAGFDYKSDSLQNTNELYAVFVGLTDGFA 236

DB 164 -----ETGKPVTLKDVFGAYSMDVITSTSGVWIDSLNPD----- 200

QY 237 PTLDSFALM-WDFV-PYFRTMK-RRHEPLTQGLAV---SREVGLMECKQKQAVLGS 290

DB 201 PFVENTKLLRFDFLDPFLSITFFPIILEVLNISI:FPREVTVFLAKSVKR-----M 255

QY 291 SQAQVKKDVQGRDILSLVRAANTAAALPESQKLSDEEVLQATSNLLFAGYETSITVLT 350

DB 256 KESRIEDTQKRVDFLQMLDWSQNSKETESHKALSDLELVAQSIIFAGYETSITVLT 315

QY 351 MFHRLSEDKAVQDKLREICQIDTDMPTLDELNALPYLEAFVKESLRLDPPSPYANRE 408

DB 316 INVELATHPDVQQLQEEBIDAVLENKAPPTVTVLQMEYLDVMVNETLRFPIAMRLERV 375

QY 409 CLKDEDFIPLAEFVIGRDSGVINEVRITKGTMTVMLPLFNINRSKFIYGEDAEFEPSWL 468

DB 376 CKQVE-----INGIFIPKGVVWVMPYSALHHPD-KYWPPEKFLPERFS 419

QY 469 EVDVTDLSNS-TEAPYHQASISCPACFGWRFAVEMKAFLEVTLLRRVQFEP 520

DB 420 KKNKNIDPIYITPFG-----SGPRNCIGRFAIMMKLAIIRVLQNFSPKP 466

RESULT 9

CP3B_MACFA

ID CP3B_MACFA STANDARD; PRT; 503 AA.

AC F3288; P25231;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome P450 3A8 (EC 1.14.14.1) (CYP11A8) (P450-MKNF2) (P-450-MK2)

DE GN

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=93129612; PubMed=1282830;

RA Komori M., Kikuchi O., Sakuma T., Funaki J., Kitada M., Kametaki T.;

RT Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity

RT of the primary sequences to human cytochromes P-450.;

RL Biochim. Biophys. Acta 1171:141-146(1992).

RN [2]

RP SEQUENCE OF 1-22.

RC TISSUE=Liver;

RX MEDLINE=89287352; PubMed=2500151;

RA Ohta K., Kitada M., Hashizume T., Komori M., Ohi H., Kametaki T.;

RT Purification of cytochrome P-450 from polychlorinated biphenyl-

RT treated crab-eating monkeys: high homology to a form of human

RT cytochrome P-450.;

RL Biochim. Biophys. Acta 996:142-145(1989).

CC -1- FUNCTION: Catalyzes nifedipine and nifedipine oxidations.

CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +

CC oxidized flavoprotein + H(2)O.

CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.

CC -1- INDUCTION: By polychlorinated biphenyl (PCB).

CC -1- SIMILARITY: Belongs to the cytochrome P450 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=ddv; TISSUE=Liver;
RX MEDLINE=9223116; PubMed=1339292;
RA Yanagimoto T., Itoh S., Muller-Eberhard D., Kametani T.;
RT "Mouse liver cytochrome P-450 (P-450I1A1): its cDNA cloning and
RL inducibility by dexamethasone.";
RN Biochim. Biophys. Acta 1130:329-332(1992).
RP [2];
RC SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Souffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- FUNCTION: Catalyzes erythromycin N-demethylation, nifedipine
CC oxidation and testosterone 6 beta-hydroxylation.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: Highly expressed in liver.
CC -!- INDUCTION: By dexamethasone.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X60452; CAA42981.1; -.
CC EMBL; BC010528; AAH10528.1; -.
CC PIR; S22334; A60564.
CC HSP; P14779; LJPZ.
CC MGI; MGI:88609; Cyp3a11.
CC InterPro; IPR001128; Cytochrome_P450.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01689; EP4501CYP3A.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 504 AA; 57854 MW; E369AF71CE23F180 CRC64;
Query Match 13.8%; Score 394; DB 1; Length 504;
Best Local Similarity 25.6%; Pred. No. 1.2e-15;
Matches 144; Conservative 103; Mismatches 215; Indels 100; Gaps 23;

```

```

CC      predominantly female-specific expression in livers.
CC      -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AB033414; BAA95951.1; -.
CC      DR MGD; MGI:1858451; Cy3a41.
CC      DR InterPro; IPR001128; Cytochrome P450.
CC      DR InterPro; IPR008072; EP450_CYP3A.
CC      DR Pfam; PF00067; P450; 1.
CC      DR PRINTS; PR01689; EP450IICYP3A.
CC      DR PRINTS; PR00385; P450.
CC      DR PROSITE; PS00086; CYTOCHROME P450; 1.
CC      KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC      KW Microsome; Endoplasmic reticulum.
CC      FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC      SEQUENCE 504 AA; 57959 MW; 2BC645B6E9CD48A CRC64;
CC
CC      Query Match      13.7%; Score 391; DB 1; Length 504;
CC      Best Local Similarity 26.2%; Pred. No. 1.8e-15;
CC      Matches 148; Conservative 100; Mismatches 212; Indels 104; Gaps 23;
CC
CC      QY      7 LTGALGLAAPSWSIASIAFFSLY-LAPERSSLYNLQG-PNHT-----NYFTGNFLDIL 55
CC      Db      3 LFSALSLSLTWLLAILVLVLYRYGTRTHGLFKKQIGIPGTPLPFLGTVLNYKGLW---- 58
CC
CC      QY      56 SARTGEHAKYREKYSTL-RFAGIAGAPVLNSTDPKFNHYM-KEAY-----DYPKP 106
CC      Db      59 -----KFDMECYEKYKTNGLFDG--QMPLFVITDPEMIKNVLKVCPSVFNRRFGVP 111
CC
CC      QY      107 GMAARVLRTAGDVVTAGEAHKRRIMISLSAQAVKSNVPFLEKGNELADKWMED 166
CC      Db      112 GIMSKAISI-----SKDEWKRYRALLSTFTSGKLKEMFPPIQYDILVKYLMQE 163
CC
CC      QY      167 AAEKDMVAGSAGEKKATRLTEGVVDKDWGRATLDVMAIAGFDYKDSIQNTNELYV 226
CC      Db      164 -AEKG-----KPTMKDVLGAYSIDVITSFGVVDLSLNNPDPPE 205
CC
CC      QY      227 AFVGL--TDGAFATLSPFAIMWDFV-PYFRTWKRHEIPLTQGLAVSRVGIEMEQK 283
CC      Db      206 KAKGILRVDFDLV--FSVLPFLTPVYEML-----NICMFPKDSIEFFPK-- 251
CC
CC      QY      284 QAVLGASDQAVKDKVQGRDILSLVRANIAANLPESOK-LSDEVLQAQISNLLFAGYE 342
CC      Db      252 --FVNRKESRLSKQKRVDFLQLMNHNNSKDKSHKALSDMEITAQSVIFIFAGYE 309
CC
CC      QY      343 TSSTVLFWPHRLSEDAKVDKLRREICQI--DTDMPTLDELNALPYLEAFVKESLRIDP 400
CC      Db      310 TTSSTLSFTLYCLATHPDIOKKLQEBIDETLPNKAPPTYDTVMEMEYLDMLVNETRLYP 369
CC
CC      QY      401 PSPANKECLKDEDFPLAEFVIGRSGSVINEVRLTKGMVWMLPLFNINRSKFIYGEDAE 460
CC      Db      370 IGRNLEFCKDVE-----LNGVTPKGTVMIPSYALHHPDQHPPE-PE 413
CC
CC      QY      461 EFRPERLWDTVDSLSNIEAPYGHQASFGISGPRACFGWRFPAVAKMKAFLFTVLRVQPEP 520
CC      Db      414 EFQPERFSKNGKSID----PYLYM-PFGIGPRNCIGMRFAPMTWTKALTKVMQNFSPQ 468
CC
CC      QY      521 IISHPEVHEHTLISRPVIGREK 544
CC      Db      469 C-----QETQPLKLSROGLLOPEK 488
CC
CC      RESULT 12
CC      CP37 HUMAN
CC      ID -CP37 HUMAN
CC      AC P24462;

```

```

DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Cytochrome P450 3A7 (EC 1.14.14.1) (CyP3A7) (P450-HFLA).
GN      CYP3A7.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RN      SEQUENCE FROM N.A.
RC      TISSUE=Fetal liver;
RX      MEDLINE=89255154; PubMed=2722762;
RA      Komori M., Nishio K., Ohi H., Kitada M., Kamataki T.;
RT      "Molecular cloning and sequence analysis of cDNA containing the
RT      entire coding region for human fetal liver cytochrome P-450.";
RL      J. Biochem. 105:161-163(1999).
RN      [2]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=21163842; PubMed=11266076;
RA      Gellner K., Eiselt R., Huster E., Arnold H., Koch I., Haberl M.,
RA      Deglmann C.J., Burk O., Buntefuss D., Escher S., Bishop C.,
RA      Koebe H.-G., Brinkmann U., Klenk H.-P., Kleine K., Meyer U.A.,
RA      Wojnowski L.;
RT      "Genomic organization of the human CYP3A locus: identification of a
RT      new, inducible CYP3A gene.";
RL      Pharmacogenetics 11:111-121(2001).
CC      -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC      monooxygenases. In liver microsomes, this enzyme is involved in an
CC      NADPH-dependent electron transport pathway. It oxidizes a variety
CC      of structurally unrelated compounds, including steroids, fatty
CC      acids, and xenobiotics.
CC      -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC      oxidized flavoprotein + H(2)O.
CC      -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC      -1- INDUCTION: P450 can be induced to high levels in liver and other
CC      tissues by various foreign compounds, including drugs, pesticides,
CC      and carcinogens.
CC      -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC      -1- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC      NOTE=CYP3A7 alleles;
CC      WWW="http://www.imm.ki.se/CYPalleles/cyp3a7.htm".
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; D00408; BAA00310.1; -.
CC      DR EMBL; AF280107; AAG32289.1; -.
CC      DR PIR; JX0062; JX0062.
CC      DR HSSP; P14779; LJPZ.
CC      DR Genew; HGNC:2640; CYP3A7.
CC      DR MIM; 605340; -.
CC      DR GO; GO:0015034; P:cytochrome P450 activity; TAS.
CC      DR InterPro; IPR001128; Cytochrome P450.
CC      DR InterPro; IPR008072; EP450_CYP3A.
CC      DR Pfam; PF00067; P450; 1.
CC      DR PRINTS; PR01689; EP450IICYP3A.
CC      DR PRINTS; PR00385; P450.
CC      DR PROSITE; PS00086; CYTOCHROME P450; 1.
CC      KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC      KW Microsome; Endoplasmic reticulum.
CC      FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC      SEQUENCE 503 AA; 57470 MW; 087CCED9BAC314C CRC64;
CC
CC      Query Match      13.5%; Score 385.5; DB 1; Length 503;
CC      Best Local Similarity 29.0%; Pred. No. 3.8e-15;
CC      Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

```

QY 13 LAAPSW--ASTAPSLYL-APRSSLNVLQ-PMHTNY-FTGNFLDILSARTG-----E 61
DQ 7 LAVETWLLAVSLILLYGTRTHGLFKKIGIPGTPPLPLGN---ALSFRKGYWTFDME 63
QY 62 EHAKYREKYSTLRFAGTAGAPVNLSDPKVFNHWM-KEAYD-----YKPPGMAARVL 113
DQ 64 CYKXKRWG-----YDCQPMALITDPDMIXTLVKECYSVFTNRFPFGVGMNAI 118
QY 114 RIATGQGVVTAAGBAHKRRIMIPISAOAKSMVPIFEKGMELVDKMDAEAKDWA 173
DQ 119 SI-----AEDBEWKRIRLSFTTSKGLKEMVPIIAQYQDVLVRLRREA----- 164
QY 174 VGESAGEKKAATLET-EGVDVQKDWGRATLDVWALAGFDYKSDLNKNTNLYVAVVGLT 232
DQ 165 -----BTGKVLKLVFGAYSMDVITSFGVSIOSLNNPQD-----PVENT 207
QY 233 D---GFAPTLDSFKAIMWDFVFPYRTMKRRHEIPLTQGLAVS---RRVGIELMEQKKQAV 286
DQ 208 KKLIRFNP-LDPFVLSIKVF-PFL-----TPILEALNITVPRKVISFLTKSVKQIK 257
QY 287 LGSASDAQVKKQVQGR--DILSLVVRANTAANLPESQKLSDEEVLQAI-SNLLFQGYETS 344
DQ 258 EGRL-----KETQKHVRFLQMLDSQNSKQSETHKALSDELMAQSIIFIFAGYETT 310
QY 345 STVLTWMPHRLSEKAVQDKLEBICQIDTDM-----PTLDELNALPYLEAFVYKESLRLD 399
DQ 311 SSVLSFIIYELATHPDVQCKVQKE--IDTVLPNKAPPTVDTVLQLEYLDMVNVNLTRLF 367
QY 400 PPSVYANRECLKDEDFIPLABPVTGROGVNVEVITKGTWMLPFLFNINRSFVIGEDA 459
DQ 368 PYAMELERVCKDVE-----INGMFIKGVVVMIPSVLVLHDP-KWTTP 411
QY 460 BEFRPERWLEDVDSLNS-IEAPYGHQASF-SGPRACFGWRFAVAKFLFVTLRRVQF 518
DQ 412 EKFLPERFSKKKNIDYIITPG-----SGPRNCIGRFAVNVNKLALVRVQLNFSF 465
QY 519 EP 520
DQ 466 KP 467
RESULT 13
CP35 HUMAN
ID CP35 HUMAN STANDARD; PRT; 502 AA.
AC P20815; Q9HB56;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome P450 3A5 (EC 1.4.14.1) (CYP11A5) (P450-PCN3).
GN CYP3A5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89278095; PubMed=2732228;
RA Aoyama T., Yamano S., Waxman D.J., Lapenson D.P., Meyer U.A.,
RA Fischer V., Tyndale R., Inaba T., Kalcow W., Gelboin H.V.,
RA Gonzalez F.J.;
RT "Cytochrome P-450 hPCN3, a novel cytochrome P-450 IITA gene product
RT that is differentially expressed in adult human liver. cDNA and
RT deduced amino acid sequence and distinct specificities of cDNA-
RT expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and
RT cyclosporine.";
RL J. Biol. Chem. 264:10388-10395 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Colon;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalious D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE OF 1-106 FROM N.A.
RX MEDLINE=21163842; PubMed=11266076;
RA Gellner K., Eisel R., Huestert E., Arnold H., Koch I., Haberl M.,
RA Deglmann C.J., Burk O., Buntfuss D., Escher S., Bishop C.,
RA Koebel H.-G., Brinkmann U., Klenk H.-P., Klein K., Meyer U.A.,
RA Wojnowski L.;
RT "Genomic organization of the human CYP3A locus: identification of a
RT new, inducible CYP3A gene.";
RL Pharmacogenetics 11:111-121 (2001).
RN [4]
RP SEQUENCE OF 1-24 FROM N.A.
RX MEDLINE=96192836; PubMed=8569713;
RA Schuetz J.D., Schuetz E.G., Thottassery J.V., Guzelian P.S., Strom S.,
RA Sun D.;
RT "Identification of a novel dexamethasone responsive enhancer in the
RT human CYP3A5 gene and its activation in human and rat liver cells.";
RL Mol. Pharmacol. 49:63-72 (1996).
RN [5]
RP VARIANT CYP3A5*2
RX MEDLINE=96192071; PubMed=8619878;
RA Jounaidi Y., Hyrailles V., Gervot L., Maurel P.;
RT "Detection of CYP3A5 allelic variant: a candidate for the polymorphic
RT expression of the protein?";
RL Biochem. Biophys. Res. Commun. 221:466-470 (1996).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2O).
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP3A5 alleles;
CC WWW="http://www.imm.ki.se/CYPalleles/cyp3a5.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; J04813; AAA02993.1; -;
CC EMBL; BC033862; AAH33862.1; -;
CC EMBL; AF280107; AAG32288.1; -;
CC EMBL; L35912; AB000083.1; -;
CC EMBL; A34101; A34101.
CC HSP; P14779; 1JFZ.
CC Genew; HGNC:2638; CYP3A5.

```

DR MIM; 605325; -.
DR GO; GO:0005792; C:mitochondrion; TAS.
DR GO; GO:0015034; F:cytochrome P450 activity; TAS.
DR GO; GO:0004497; F:monooxygenase activity; TAS.
DR GO; GO:0008202; P:steroid metabolism; TAS.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Polymorphism.
FT METAL 441 441 IRON (HEME AXIAL LIGAND).
FT VARIANT 398 398 T -> N (in allele CYP3A5*2).
FT /FTID=VAR_008365.
SQ SEQUENCE 502 AA; 57108 MW; D5A2302E2633E717 CRC64;
Query Match 13.5%; Score 385; DB 1; Length 502;
Best Local Similarity 28.1%; Pred. No. 4.1e-15;
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;
AC P79102; STANDARD; PRT; 507 AA.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A28 (BC 1.14.14.1) (CYP11A28).
GN CYP3A28.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;

```

```

[1]
RN RP
RC STRAIN=Hereford; Tissue=Liver;
RA Natsuhori M., van Raak M., Lichtenberg M., Kleij L., ten Berge D.,
RA Zweers-Zeilmaker W.M., de Groene E.M., van Miert A.S.J.P.A.,
RA Witkamp R.F., Horbach G.J.M.J.;
RT "Isolation of a bovine full length cytochrome P450 (CYP3A) cDNA
RT sequence and its functional expression in V79 cells.";
RL Environ. Toxicol. Pharmacol. 3:17-24(1997).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y10214; CAAT1266.1; -.
DR HSSP; P14779; IJFZ.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT SEQUENCE 507 AA; 58152 MW; FDFD8465CC96DF57 CRC64;
Query Match 13.5%; Score 385; DB 1; Length 507;
Best Local Similarity 25.5%; Pred. No. 4.1e-15;
Matches 140; Conservative 94; Mismatches 132; Indels 122; Gaps 22;
QY 38 LQGNHTNYFTGNFLDILSARTGEEHAKYR-----EKYGSTLRPAGIAGAPVLNS 87
DB 38 IGPRLPVYF-----GSTMAVHKGIPEDFNQCFKFKYGMWGFYE-GRQPLAI 84
QY 88 TDPKVENHYM-KEAYD-----YKPGMAARVLRIATGDGVVTAEGEAHRRHRIMPS 139
DB 85 TDPDIKTIVLKECVSVFTNRIRIFGPMIMKVALSLW-----DEQWKIETLLSPA 136
QY 140 LSAQAVKSMVPFLEKGMELVDKMDAAEKMAVGESAGEKKAATRLTETEGVDVQWVGR 199
DB 137 FTSGKLKEMFPIIGQYGDMLVRLNRKE-AEKG-----NPNVNMKDMFGA 178
QY 200 AILDWALAGPDYKSDSLQNKTNELYAVFVGLTGDGAPTLDSF--KAIMDEV-PYERTM 256
DB 179 YSMDEVITGTAFGNIDSL--NNPHDPFVEHSHKNLRLFRP-FDPFILSLILFFINPVEIL 236
QY 257 KRRHEIPLTQGLAVSRVRGIELMEQKQAVLGSASDAQVKKQVQGRDIISLVANIAA 316
DB 237 -----NITLFPKSVDFTKSVKKI---KESRLTDKQNRVLDLLQLMINSQNSK 282
QY 317 NLPESOKLSDEVLQAISNLLPAGYETSTVLTWTFHRLSEDKAVQDKLREIICQITDM 376
DB 283 EIDNHKALSDIELVAQSTIFIFGGYETSTLSFIYELTTHPHVQKQKE---IDATF 339
QY 377 -----PTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAEPVIGRDSVIN 431

```


Db 340 PNKAPFYDALVQMEYLDVMVNETLRFIAGRLSERVCKOVE-----IH 384
 QY 432 EVRITKGTVMVLEPLFNINRSKFTYGDABEERPERMLEDVTDLSNS-IEAPYGHQASFIS 490
 Db 385 GVPIPKGTTVLVPLFLVHNHPPELWPE-PEEPRPERFSKNNKDSINPVVLPFG-----T 437
 QY 491 GPRACGWRFAVAKAFLEFVTLRRVQPEPI-----ISHREYEHITLISRP 537
 Db 438 GPRNCLGMRFAIMNVIKALVRILONFSPKCKETQ:PLKLYTQGLTQBEQPVILKVW--P 495
 QY 538 RIVGREKE 545
 Db 496 RGLGPQVE 503

RESULT 15
 ID -CP3L CALJA STANDARD; PRT; 503 AA.
 AC O18953;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 3A21 (EC 1.14.14.1) (CYP11A21) (P450 CM3A-10).
 GN CYP3A21.
 OS Callithrix jacchus (Common marmoset).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
 OC Callithrix.
 OC Callithrix.
 OX NCBI_TaxID=94483;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TIGR=Liver;
 RX MEDLINE=97223367; PubMed=9056237;
 RA Igarashi T., Sakuma T., Isogaki M., Nagata R., Kamataki T.;
 RT "Marmoset liver cytochrome P450s: study for expression and molecular
 RL cloning of their cDNAs."
 RL Arch. Biochem. Biophys. 339:85-91(1997).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other
 CC tissues by various foreign compounds, including drugs, pesticides,
 CC and carcinogens.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D31921; BAA22156.1; -;
 DR HSSP; P14779; 1JPZ.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008072; EP450_CYP3A.
 DR Pfam; PF00667; P450; 1.
 DR PRINTS; PRO1689; EP450ICYP3A.
 DR PRINTS; PRO0385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 FT Microsome; Endoplasmic reticulum.
 KW METAL. 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 503 AA; 57564 MW; 9BFD42ID72C76D6A CRC64;

Query Match 13.2%; Score 377.5; DB 1; Length 503;
 Best Local Similarity 27.1%; Pred. No. 1.1e-14;

Matches 148; Conservative 90; Mismatches 184; Indels 125; Gaps 26;
 QY 13 LAAPSW--ASIAFFSLYADPRRS--SLYNLOQ--PHHTNY--FTGNFLDILSARTG-----E 61
 Db 7 LAVETWLLAVSLVLYLYGTHSHGLFKKLGIPGPTPLPFLGT---VLYYRQGFWEKFDME 63
 QY 62 EHAKYREKYGSLRREAGIAGAPVLNSTDPKVFNVHM--KEAYD-----YPKPGVAARVL 113
 Db 64 CYKKYKGMWGI---YDG--RQPVLAITDPNIIKIVLKECYSVFTNRPRFPVPGVFMKSAI 118
 QY 114 RIATDGVVTAEGEAHGRHRRIMIPSLSAQAVKSNVPIFEKGMELVDKMDAABKDMA 173
 Db 119 STAQDD-----EWKQIRSLSPFTFTSGKLEKEMVPIIAQYGEVLVRNLARE--AEKG-- 167
 QY 174 VGESAGEKKALETETEGVDVKDWGRATLDVNALAGEDYKSDSLQNKTNELVYAFVGLTD 233
 Db 168 -----KPINMKDIFGAYSMDVITGTSFGVWIDSLNNPD----- 201
 QY 234 GFAPTLDSFKAIM--WDFV--PYFRVTKRRHEI--PLTQGLAVSR-----RVGIELMEQ 281
 Db 202 ---PFVESTKLLRFDLDPDFLPPFLSITIPFPLTILEALNISMFPRDSTSLRKGIKRIKE 258
 QY 282 KKOAVLGSASDAQVKDVQGR--DILSLVYRANITANLPESQKLSDBEVLQAQISNLLFA 339
 Db 259 SR-----LKOYTHKRVDFLOQIMDSQNKETESDKALSDELVAQSIIFIFA 305
 QY 340 GYETSSVTILTMFHRILSEDKAVQDKLREEICQIDTDM-----TLDELNALPYLEAFVKE 394
 Db 306 GYETTSSTLSFIMYELATHPDVQKQLOBE---IDAVLPNKAPATYDVLQMEYLDVMVNE 362
 QY 395 SLRLDPPSPYANRECLKDEDFIPLAEFVIGRDSVINEVRITKGTVMVLEPLFNINRSKF1 454
 Db 363 TLRFLFLAVRLSERVCKOVE-----INGVFIPKGVVVMIPSYALHYDP-K 406
 QY 455 YGSDAEERPERMLEDVTDLSNS-IEAPYGHQASFISGPRACFGWRFAVAKAFLEFVTL 513
 Db 407 YWTEPEKFLPERFSKNNKDNIDPIYITPFG-----TGPNCIGWRFALMNMKALIRVL 460
 QY 514 RRVQFEP 520
 Db 461 QNFSPKP 467

Search completed: April 2, 2004, 13:59:09
 Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17 seconds
(without alignments)
3151.687 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLLTGALGLAFAFWAS.....RIVGREKEGYQWRLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

FIR_78.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	419	14.7	593	2 F86441	probable cytochrom
2	406	14.2	503	2 JC4702	cytochrome P450 3A
3	404	14.2	503	2 S14275	steroid 6beta-mono
4	402	14.1	501	2 A34236	cytochrome P450 3A
5	400	14.0	526	2 T02191	cytochrome P450 ho
6	399	14.0	504	2 A29410	cytochrome P450, g
7	398.5	14.0	503	2 A29815	cytochrome P450 3A
8	398	14.0	503	2 S50211	cytochrome P450 3A
9	395.5	13.9	503	2 S28168	nifedipine oxidase
10	394	13.8	504	2 A60564	cytochrome P450 3A
11	385.5	13.5	503	2 JX0062	cytochrome P450 3A
12	385	13.5	502	1 A34101	cytochrome P450 3A
13	379	13.3	501	2 A29487	cytochrome P450 3A
14	370	13.0	504	2 A22631	cytochrome P450 3A
15	368	12.9	504	2 S50832	cytochrome P450 3A
16	365	12.8	520	2 H84663	probable cytochrom
17	363.5	12.7	511	2 T00864	probable cytochrom
18	362.5	12.7	523	2 B96662	probable cytochrom
19	359.5	12.6	511	2 S66472	cytochrome P450 4B
20	357	12.5	502	2 JX0334	cytochrome P450 3A
21	355.5	12.5	512	2 A96695	hypothetical prote
22	350.5	12.3	1054	1 A69975	NADPH-ferritinemo
23	350	12.3	526	2 JC4533	cytochrome P450 4F
24	349.5	12.3	511	1 O4HUB1	cytochrome P450 4B
25	349.5	12.3	524	2 S29723	cytochrome P450 4F
26	346.5	12.1	511	1 B40164	cytochrome P450 4B
27	346.5	12.1	524	2 JC7594	cytochrome P450 en
28	346.5	12.1	524	2 JC7598	cytochrome P450 en
29	346	12.1	537	2 T02450	probable cytochrom

ALIGNMENTS

RESULT 1

F86441

Probable cytochrome P450 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001

C:Accession: F86441

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;

ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:1130712

A:Accession: F86441

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-593 <STO>

A:Cross-references: GB:AE005172; NID:g11136728; PIDN:AAG31309.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology

C:Keywords: heme; iron; metalloprotein

F:514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.7%; Score 419; DB 2; Length 593;

Best local similarity 26.4%; Pred. No. 1.2e-19;

Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NFDILSARTG--EEHAKYREKYGST-----LRPAGI-----AGAPVLN 86

DB 93 NVLDFMDTGSDDYKPKVPEAKGSIQAVNEAFIPLVEFLTYGGIFRLTFGPKSFLI 152

QY 87 STPKVFNHWKKE-NDYDKPGMAARVLRIATGDGVTAEAGAHKRRHRIMIPSLSAQAV 145

DB 153 VSDPSIAKHILKONAKAYSK-GILAEILDFVMKGKLIIPADGEIWRRRRAIVPALHOKVY 211

QY 146 KSNVPFIFLEKGMVLDRMMDEAAEKDMAYGESAGEKATRLTEGVVDKXWVGRATLDVM 205

DB 212 AAMISLFGEASDRLCQKL--DMA-----ALKGEVEVMSLSRLTLDI 253

QY 206 ALAGFDYKSLQNKTNELVAFVGTGDFAPLDSFKAIMWDFVFPFTRKRRHEIPIT 265

DB 254 GKAVFNDFDPSLNDTGVIEAVTVLREAEKRSVSDIP--VMD-IPFWKDS-----FRQ 305

QY 266 QGLAVSRVVGIELMEKKQAVLGASDQAVDKKDVQ-----GRD--ILSLVLRANIAA 316

DB 306 RKVATSLKLNILDD-----LIATCKRVVEBELQFHEEYMNDRDPSILHFL----- 353

hypothetical prote
cytochrome P450 ho
leukotriene B4 ome
cytochrome P450 3A
probable cytochrom
probable cytochrom
probable cytochrom
cytochrome P450 3A
leukotriene-B4 20-
probable cytochrom
cytochrome P450 4F
hypothetical prote
probable cytochrom
cytochrome P450III
hypothetical prote

RESULT 4
A34236

cytochrome P450 3A6 (version 2) - rabbit
 N:Alternate names: cytochrome P450 3C
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 28-Jul-2000
 C:Accession: A34236
 J:Portenza, C.L.; Pendurthi, U.R.; Strom, D.K.; Tukey, R.H.; Griffin, K.J.; Schwab, G.E.;
 J:Biochem. 264, 16222-16228, 1989
 A:Title: Regulation of the rabbit cytochrome P-450 3C gene. Age-dependent expression and
 A:Reference number: A34236; MUID:89380226; PMID:2777787
 A:Accession: A34236
 A:Molecule type: mRNA
 A:Residues: 1-501 <POT>
 A:Cross-references: GB:J05034; NID:g164829; PIDN:AAA31178.1; PID:g164830
 C:Genetics:
 A:Gene: CYP3A6
 C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;
 F:300-462/Domain: cytochrome P450 homology <P45>
 F:440/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.1%; Score 402; DB 2; Length 501;
 Best Local Similarity 27.8%; Pred. No. 1.2e-18;
 Matches 156; Conservative 88; Mismatches 200; Indels 118; Gaps 24;

QY 19 ASIAFFSLYLAPRRSSLYNLOQPNHTNY-FTGNFLDILSARTG--BEHAKYREKYGSTLR 75
 DB 15 ASLVLLYLYGTSHTGLFKKMGIPGPTPLPFIPTG---ILEYRKGIDWDFIECRKKYG--K 68

QY 76 FAGI--AGAPVNLSTDPKVFNVHM-KEAY-----DYPKPGMAARVLRATGCGVVTAE 125
 DB 69 MWGLFDGQPLVITDPPMIKTLVLCYVSFTNRRSGFVGFVGMKKAVSI-----SE 120

QY 126 GEAKHRRIRIPSLSAQAQKSMVPIFEKGMELVDKMDAEAKDMVAVGESAGEKKATR 185
 DB 121 DEDWKRVTLSPTFTSKLKEMLPITAOYGDVLV-KNLRQEAEG-----165

QY 186 LETEGVDYKDWGRATLDVMAALGDFYKSDSLQNKTN-----LYAVFGLTDFGAP 237
 DB 166 ---KEVDLKEIFGAYSMVDVITGTSFGVNIDSLRNPQDPFVKNVRLRLKFSF-----FDP 216

QY 238 TLDSEKAIWMDVFPVFRTRKRRHRIPLTQGLAVSRVVGIELMEQKQAVLSASDAQVK 297
 DB 217 LLSL-----TLFPL-----TPFEALHIS-MPFKQVMDFLKTSVEKIKDDRLKK 262

QY 298 KDVGQRDILSLVRANAIANLPEOKLSDEEVLQAISNLLFAGYETSSVTLTWMFRLSE 357
 DB 263 QK-RRVDFQLMINSQNSKEIDSHKALDDIEVVAQSIILFAGYETSSVTLSPIMHLLAT 321

QY 358 DRVQDKLREEICQIDTDP-----TLDELNALPYLEAFYKESLRIDPPSPYANRECLKD 412
 DB 322 HPDVQOKLQEE---IDTLNPKELATYDTLVKMEYLDVMVNETURLYPIAGRLERVCKXD 378

QY 413 EDFIPLAEFVIGRDSVINEVRITKGTVMPLFENIRSKFTYGEDAEFPRPRLWLEVDYT 472
 DB 379 VD-----INGFIPKGTIVMPTVALHRDPQHWT-EDEFRPERFSKKN 422

QY 473 DLSNS-IEAPYGHQASFTSGPRACGFWFAVAKAFVTLRRVQPE-----PI---521
 DB 423 DNINPYIYHPFG-----AGPRNCILGMRPALMNKIALVRLMQNFSPLCKETQVPLKLG 476

QY 522 ---ISHPEVEHETLISPRIV 540
 DB 477 KQGLLOQPEKPIVJKVVSRDGII 498

RESULT 5
 T02191
 cytochrome P450 homolog F14M4.21 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02191; D84909
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rot

submitted to the EMBL Data Library, September 1998
 A:Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.
 A:Reference number: Z14609
 A:Accession: T02191
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-526 <ROU>
 A:Cross-references: EMBL:AC004411; NID:g3522932; PID:g3522945
 A:Experimental source: Cultivar Columbia
 A:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: D84909
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-526 <STO>
 A:Cross-references: GB:AE002093; NID:g3522945; PIDN:AAC34227.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g48960; F14M4.21
 A:Map position: 2
 A:Introns: 93/1; 168/3; 254/2; 375/3
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
 C:Keywords: heme; iron; metalloprotein
 F:322-493/Domain: cytochrome P450 homology <P45>
 F:471/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 400; DB 2; Length 526;
 Best Local Similarity 23.6%; Pred. No. 1.7e-18;
 Matches 140; Conservative 109; Mismatches 231; Indels 114; Gaps 21;

QY 1 MFIVLLTGLGAAAFSWASIAFSLYLAPRRSSLYNLOQPNHTNYFTGNF-----51
 DB 9 LIVILLI---IGLRIFKAFVILVWHPVLRLLKNOGIGSPNY-RIFYGNLSEIKOMKE 64

QY 52 --DLISARTGE-----EHAKYREKYSTLRFAGIAGAPVNLSTDPKVFNVHWEAYD 102
 DB 65 SHLSILDPSSNDIFPRILPHYQKMSQYGETFLYWN-GTEPRICISDPELAKTMLSNKLIG 123

QY 103 -YPPGMAARVLRATGCGVVTAEAGHAKHRRIRIPSLSAQAQKSMVPIFEKGMELVD 161
 DB 124 FVTSKARPEAKVLGSKGLVFIEGADWVRHRLINPAPSIDRLKIMTVMVDCITL----179

QY 162 KOMEDAAEKDMVAVGESAGEKATRLTEGVVDKDWVG-----RATLDVMAALGDFYKSDSL 217
 DB 180 KMLEE-----WRKESTKEETHPKIKKENNEEFORLTADIIATSAFG-----221

QY 218 QNKTNELVAVFGLTDFGAPFLDSFKAIWMDVFPVFRTRKRRHRIPLTQGLAVSRVVGIE 277
 DB 222 -----SSYVEGIEVFRSQM-ELKRCVTTSLNQSIVPGTQYLPFSNIRW 265

QY 278 LMEOKQAVLSASDAQVDKQVQGRDILSLVRANAIANLPEOKLSDEEVLQAISNLL 337
 DB 266 KLERKMDNSIKRIISRLSKSDYGDGLLGLLKAYNTEG--KERKMSIEEIIHECRTEFF 323

QY 338 FAGYETSSVTLTWMFRLSEDKAVQDKLREEI---C-QIDTMDPLDELNALP-----YL 388
 DB 324 FGGHETTSNLLAWTMTMLSLHQDKQKRLKEEIPKECGKKTDPDSETFSLKUKPVVFLQM 383

QY 389 EAFVKESLRIDPPSPYANRECLKDEDFIPLAEFVIGRDSV---INEVRITKGTVMVLEL 445
 DB 384 NWVIMESLRLYGP-----VSEALAREASVNIKGLDLBPKGTTVWIFL 425

QY 446 FNINRSFTYGEDAEFPRPRLWLEVDYDLSNLSNTEAPYGHQASFTSGPRACGFWFAVAK 505
 DB 426 LKMHSDKTLWGSADKFNPMRFANGVSRANHPNA-----LLAFSVGPRACIQGNFWIEA 481

QY 506 KAFVLTLLRRVQPEPIIISHPEYEH--ITLLISPRIVGRKEGYQMRLOVKVPE 557
 DB 482 KTVLTMLQORRFFSLQCD--EYKHTPVDNVTIQPQ-----YGLPVMVLQPLE 525

RESULT 6

A29410
cytochrome P450, glucocorticoid-inducible, hepatic - human
N:Alternate names: cytochrome P450 3A3; cytochrome P450 H1p
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1989 #sequence revision 31-Mar-1989 #text_change 28-Jul-2000
C:Accession: A29410; A25170; A25450
R:MoLow, D.T.; Schuetz, E.G.; Wrighton, S.A.; Watkins, P.B.; Mendez-Picon, R.; Molowa, D.T.; Schuetz, E.G.; Wrighton, S.A.; Schuetz, E.G.; Mendez-Picon, G.
Proc. Natl. Acad. Sci. U.S.A. 83, 6310-6314, 1986
A:Title: Complete cDNA sequence of a cytochrome P-450 inducible by glucocorticoids in hu
man liver.
A:Reference number: A25170; MUID:85298342; PMID:3898085
A:Accession: A25170
A:Molecule type: protein
A:Residues: 2-21 <WAT>
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; liver; metalloprotein; monooxy
F;303-465/Domain: Cytochrome P450 homology <P45>
F;443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 399; DB 2; Length 504;
Best Local Similarity 28.4%; Pred. No. 1.9e-18;
Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;

Qy 13 LAAPSW--ASIAFFSLYLAPRS-SLYNLQG-PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAMETLLLAVALSVLLYLYGTHSHGLFKLGGPGTLPFLGN---ILSYHKGCFMDME 63
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYD-----YKPGMAARVL 113
Db 64 CHKKYKQWGF---YDQ--QQVLAITDPDMIKLVLEKCYSVFTNRRPFGVGMKSAI 118
Qy 114 RIATGDGVTAEGEAHQRHRRIMPSLSAQAVKSMVPIFLKGMELVDKMWEDAAEKDMA 173
Db 119 SI-----AEDEEWKRLSLSPFTSGKLKEMVPIIAQYGDVVLVNRRE----- 163
Qy 174 VGSAGEKATRLTEGVDVQWVGVRATLDVNLALAGFDYKSDSLQNKTNELYVAVGLTD 233
Db 164 -----RETGKPVTLKQVFGAYSMVDVITSSFGVNVDSLNNPD----- 201
Qy 234 GFAPTLDSFKAIM-WDFV-PYFTMK-RRHEIPLTQGL-----AVSRVVG 275
Db 202 ---PLVENTKKLRFDFLDFLDFLDFLDFLDFLDFLDFLDFLDFLDFLDFLDFLDFLDF 258
Qy 276 IELMEQKQAVLGASDAQVDKDKVQGRDILSLVLRANLAANLPESQK-LGDEEVLQIS 334
Db 259 SRLDGTQKRV-----DFQLMIDSHKNSKETSSHKALSDELVAQSI 301
Qy 335 NLLFAGVETSTVLTMMFHLSDKAVQDLREIQI---DIDMTFLDELNALPYLEAFV 392
Db 302 IFIFAGVETTSVLSFMFLMYELATHPDVQKQLOEIDAVLPKNAPPYDTVLQMEIWMV 361
Qy 393 KESIRLPPSPYANRECLDKEDFIPLAEPVIGRDSVINEVRITKGMVMLPLFNINRSK 452
Db 362 NETLRLFPANRLSERVCKQVE-----INGMFIKGMVVMIFSALHRDP 406
Qy 453 FYIGEDAEFPERLWEDVTDLSNS-IEAPYHQACFSIGSPACFQWRFAVEMKAFLEV 511
Db 407 -KYNTEBEKFLPERFSKKNKNDIPYITPFG-----SGPRNCIOMRFALMNKALTR 459
Qy 512 TLRVQPEP 520
Db 460 VLQNFSPKP 468

RESULT 7

A29815
cytochrome P450 3A4 nifedipine oxidase (EC 1.14.14.-) - human
N:Alternate names: cytochrome P450 (PCN1); cytochrome P450-HM1
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1989 #sequence revision 30-Sep-1989 #text_change 28-Jul-2000
C:Accession: A29815; A16900; A25517; A32199; PX0012; S03851
R:Gonzalez, F.J.; Schmid, B.J.; Umeno, M.; McBride, O.W.; Hardwick, J.P.; Meyer, U.A.;
DNA 7, 79-86, 1988
A:Title: Human P450PCN1: sequence, chromosome localization, and direct evidence through
A:Reference number: A29815; MUID:88195781; PMID:3267210
A:Accession: A29815
A:Molecule type: mRNA
A:Residues: 1-503 <GON>
A:Cross-references: GB:M18907; NID:G181373; PIDN:AAA35745.1; PID:G181374
R:Spurr, N.K.; Gough, A.C.; Stevenson, K.; Wolf, C.R.
Hum. Genet. 81, 171-174, 1989
A:Title: The human cytochrome P450 CYP3 locus: assignment to chromosome 7q22-qter.
A:Reference number: S16900; MUID:89108438; PMID:2563251
A:Accession: S16900
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-391, 'W', 393-503 <SP2>
A:Cross-references: EMBL:X12387; NID:G35910; PIDN:CAA30944.1; PID:G35911
R:Beaune, P.H.; Umbenhauer, D.R.; Bork, R.W.; Lloyd, R.S.; Guengerich, F.P.
Proc. Natl. Acad. Sci. U.S.A. 83, 8064-8068, 1986
A:Title: Isolation and sequence determination of a cDNA clone related to human cytochrome
P450 CYP3A4.
A:Reference number: A25517; MUID:87041402; PMID:3464943
A:Accession: A25517
A:Molecule type: mRNA
A:Residues: 1-391, 'W', 393-503 <BEA>
R:Bork, R.W.; Muto, T.; Beaune, P.H.; Srivastava, P.K.; Lloyd, R.S.; Guengerich, F.P.
J. Biol. Chem. 264, 910-919, 1989
A:Title: Characterization of mRNA species related to human liver cytochrome P-450 nifed
A:Reference number: A32199; MUID:89093163; PMID:2463251
A:Accession: A32199
A:Molecule type: protein
A:Residues: 'X', 2-9, 'XXX', 13-15 <BOR>
R:Komori, M.; Hashizume, T.; Ohi, H.; Miura, T.; Kitada, M.; Nagashima, K.; Kamataki, T.
J. Biochem. 104, 912-916, 1988
A:Title: Cytochrome P-450 in human liver microsomes: high-performance liquid chromatogr.
A:Reference number: PX0012; MUID:89214010; PMID:3243766
A:Accession: PX0012
A:Molecule type: protein
A:Residues: 'X', 2-11, 'X', 13-25 <KOM>
A:Experimental source: liver microsome
C:Genetics:
A:Gene: GDB:CYP3A4
A:Cross-references: GDB:118782
A:Map position: 7q22.1-7q22.1
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metall
F;302-464/Domain: cytochrome P450 homology <P45>
F;442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 398.5; DB 2; Length 503;
Best Local Similarity 28.7%; Pred. No. 2e-18;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

Qy 13 LAAPSW--ASIAFFSLYLAPRS-SLYNLQG-PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAMETLLLAVALSVLLYLYGTHSHGLFKLGGPGTLPFLGN---ILSYHKGCFMDME 63
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYD---PKPGMAARVLRIAT 117
Db 64 CHKKYKQWGF---YDQ--QQVLAITDPDMIKLVLEKCYSVFTNRRPFGVGMKSA- 117
Qy 118 GDGVTAEGEAHQRHRRIMPSLSAQAVKSMVPIFLKGMELVDKMWEDAAEKDMAVGES 177
Db 118 ---ISIAEDEEWKRLSLSPFTSGKLKEMVPIIAQYGDVVLVNRREA----- 164


```

QY 233 DGAPTLDSFKAIM-WDFV-PYFRTWK-RRHPIPLTQGLAVS-----RRVGIEMEQKKQAV 286
Db 202 -----PFVENTKLLRDFLDPFLSITIPFPFIPILEVINISIPPREVTSFLRKSVKRI- 256
QY 287 LGSASQAVDKVDQGR--DILSLVRANIAANLSPESOKLSDEEVLQISNLLFAGYETS 344
Db 257 -----KESRLDKTQKRVDFQLMIDSONSKETESHKALSDELELVAQSIIFIFAGYETT 310
QY 345 STVLTWVHRLSDKAVQOKLREICQIDTDM-----PTLDELNALPYLEAFVKSRLD 399
Db 311 SSVLSFIIVELATHPDVQOKLOEE-----IDTVLPNKAPPTIDTVLQWELDMVWNETLRIF 367
QY 400 PPSPYANRECLKDEDFPLAEPVIGRDSGVINEVRITGTVMVLPFLFNINRSKTIYGEDA 459
Db 368 PIAMLERVCKKDE-----INGIPIKGVVWIPSYALHDP-KYWPPEP 411
QY 460 EEPERPERWLEDVDTLSNS-IEAPYGHQASFIISGRACFGWRFAVAEMKAFILVTLRRVQF 518
Db 412 EKELPERFSKKNNDIPVIYTPFG-----SGPRNCIGMRPALNMKLAIRVLQNSF 465
QY 519 EP 520
Db 466 KP 467

RESULT 10
A60564
Cytochrome P450 3A11 - mouse
N;Alternate names: cytochrome P450 3A-UT; cytochrome P450 11A1
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Mus musculus (house mouse)
C;Date: 17-Apr-1993 #sequence_revision 17-May-1996 #text_change 28-Jul-2000
C;Accession: S22334; A60564
R;Yanagimoto, T.; Itoh, S.; Muller-Enoch, D.; Kamataki, T.
Biochim. Biophys. Acta 1130, 329-332, 1992
A;Title: Mouse liver cytochrome P-450 (P-45011A1): its cDNA cloning and inducibility b
A;Reference number: S22334; MUID:92223116; PMID:1339292
A;Accession: S22334
A;Molecule type: mRNA
A;Residues: 1-504 <YAN>
A;Cross-references: EMBL:X60452; NID:G50534; PIDN:CAA42981.1; PID:G50535
R;Bornheim, L.M.; Correia, M.A.
Mol. Pharmacol. 38, 319-326, 1990
A;Title: Selective inactivation of mouse liver cytochrome P-45011A by cannabidiol.
A;Reference number: A60564; MUID:90384441; PMID:2402224
A;Accession: A60564
A;Molecule type: protein
A;Residues: 1-17,'X',19-20,'X',22-24 <BOR>
C;Comment: This member of the cytochrome P45011A family was designated UT because it ca
lation of testosterone.
C;Genetics:
A;Gene: CYP3A11
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
P:303-465/Domain: cytochrome P450 homology <P45>
P:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.8%; Score 394; DB 2; Length 504;
Best Local Similarity 25.6%; Pred. No. 4e-18;
Matches 144; Conservative 103; Mismatches 215; Indels 100; Gaps 23;

QY 7 LTGALGAAPSWASIAFFSLY-LAPRSSI-YNLQG-----PNHTNVTGNTFLDILSAR 58
Db 3 LVSLSLTETWLLAISLVLYRGTAKHFLPKQGIPIGPKPLP-----FLGTVLNYKGL 57
QY 59 TGESEHAKYREKYGSTL-RPAGIAGAPVLNSTDPKVFHVM-KEAY-----DYPKPGWA 109
Db 58 WKFDMECYK-KYGTWGLGFDG--QTLLAVTDPETIKVNLVKECFSVFTNRDRDFGPVGIM 114
QY 110 ARVLRATGQVTAEGEAHRRIRIMISLSAQAVKSMVPIFLKGMELVDXMMEDAE 169
Db 115 SKAISIKDD-----EWKRYALLSPFTISGKUKEMFPVIEQYG-DILVKYLRQAK 165

```

```

QY 170 KDMVGESAGEKKATRLTEGVYKDWVGRATLDVVALAGFDYKSDSLQKTNELYAVFV 229
Db 166 KG-----KEVTMKDVLGAYSDMTITSTSGVNVDSLNNPDPFFVEKAK 208
QY 230 GLT--DGAPTLDSFKAIMWDFV-PYFRTMKRRHEIPLTQGLAVSRVVGIELMEQKKQAV 286
Db 209 KLLRDFDFDPL--FSVVLFPFLPVYEML-----NICMFPKDSIEFFKK--F 252
QY 287 LGSASQAVDKVDQGRDILSLVRA-NIAANLSPESOKLSDEEVLQISNLLFAGYETS 345
Db 253 VDRMKESLDSKQKRVDFLQAMNHNHNSKDKVSHKALSDEMEITAGSIIFIFAGYETS 312
QY 346 TVLTWVHRLSDKAVQOKLREECQI--DTDMPTLDELNALPYLEAFVKSRLDPPSP 403
Db 313 STLSLTHSLATHPDIOKKLODEIDEALPNKAPPTIDTVMEMEYLDVNLNETLRLYPIAN 372
QY 404 YANRECLKDEDFPLAEPVIGRDSGVINEVRITGTVMVLPFLFNINRSKIYGEDAEFR 463
Db 373 RLERYCKKDE-----LNGVYIPKGTVMIPSYALHDPQHWSE-PEEFQ 416
QY 464 PERWLEDVDTLSNS-IEAPYGHQASFIISGRACFGWRFAVAEMKAFILVTLRRVQFPEPI 522
Db 417 PERFSKENKGSIDPVYLPFG-----NGPRNCLGMRPALNMKLAITKIMQNFSPQPC- 469
QY 523 SHPEYEHITLIISPRIVGREK 544
Db 470 ---KETQIPKLRSQGLLOPEK 488

RESULT 11
JX0062
Cytochrome P450 3A7 - human
N;Alternate names: cytochrome P450 3A3; cytochrome P450 HFL33; cytochrome P450 HLP2; cy
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 28-Jul-2000
C;Accession: JX0062; EX0014; S04983; S02152
R;Komori, M.; Nishio, K.; Ohi, H.; Kitada, M.; Kamataki, T.
J. Biochem. 105, 161-163, 1989
A;Title: Molecular cloning and sequence analysis of cDNA containing the entire coding r
A;Reference number: JX0062; MUID:89255154; PMID:2722762
A;Accession: JX0062
A;Molecule type: mRNA
A;Residues: 1-503 <KOM>
A;Cross-references: GB:D00408; NID:G220148; PIDN:BAA00310.1; PID:G220149
A;Experimental source: fetal liver
R;Komori, M.; Hashizume, T.; Ohi, H.; Miura, T.; Kitada, M.; Nagashima, K.; Kamataki, T.
J. Biochem. 104, 912-916, 1988
A;Title: Cytochrome P-450 in human liver microsomes: high-performance liquid chromatogr
A;Reference number: PX0012; MUID:89214010; PMID:3243766
A;Accession: PX0014
A;Molecule type: protein
A;Residues: 1-25 <KO2>
A;Experimental source: liver microsome
R;Komori, M.; Nishio, K.; Fujitani, T.; Ohi, H.; Kitada, M.; Mima, S.; Itahashi, K.; Ka
Arch. Biochem. Biophys. 272, 219-225, 1989
A;Title: Isolation of a new human fetal liver cytochrome P450 cDNA clone: evidence for
A;Reference number: S04983; MUID:89286124; PMID:2786707
A;Accession: S04983
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 282-503 <KO3>
R;Wrighton, S.A.; Vandenberg, M.
Arch. Biochem. Biophys. 268, 144-151, 1989
A;Title: Isolation and characterization of human fetal liver cytochrome P450HLP2: a thi
A;Reference number: S02152; MUID:89104413; PMID:2492179
A;Accession: S02152
A;Molecule type: protein
A;Residues: 1-11,'X',13-25,'X',27-30 <WRI>
C;Genetics:
A;Gene: CYP3A7
A;Cross-references: GDB:134409
A;Map position: 7q22.1-7q22.1

```


C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F;302-464/Domain: cytochrome P450 homology <P45>
F;442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.5%; Score 385.5; DB 2; Length 503;
Best Local Similarity 29.08; Pred. No. 1.4e-17;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61
DB 7 LAVETWLLAVSLVLLYLGTRTHGLFKLGIQPTPLPFLGN---ALSFRKGYWTFDME 63
QY 62 EHAHYREKYGSLRFPAGIAPVLSNDPKVFNHVM-KEAYD-----YPKPGMAARVL 113
DB 64 CYKIRKVGW-----YDCCQPMALITDPNMTKVLVKECYSVFTNRRFGPVGFWKNAI 118
QY 114 RIATGDDGVVTAAGBAHRRIRIMIPSLSAQAQVSMVPIFLEKGMELVDKMDAAEKDVA 173
DB 119 SI-----ADEBEWKIRSLSPFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164
QY 174 VESAGEKKATRLT-EGVDVKDWVGRATLDVNALAGFDYKSDSLNKTNELYVAPVGLT 232
DB 165 -----ETGKPVTLKHVFGAYSDMTITSTSGVSDSLNPNQD----PPVENT 207
QY 233 D---GFAPTLDSFKAIMWDFVFPYFRTMKRRHEIPLTQGLAVS---RRVGIELMEQKKQAV 286
DB 208 KKLRFNP-LDPFVLSIKVP-PFL-----TPILEALNITVPPRKVISFLTKSVKQIK 257
QY 287 LGSASDAQVDKQVQGR--DILSLVRANTAAANLPESQKLSDEVLAAQISNLLFAGYETS 344
DB 258 EGRL-----KETQKRVDPQLQMLDSQNSKQSTHKALESLELAAQSIIFIFAGYETT 310
QY 345 STVLTWFMHRLSEDKAVQDKLREICQIDTDM-----PTLDELNALPYLEAFVKESLRID 399
DB 311 SSVLSFIYELATPDVQVQVKE---IDTVLPNKAPPTVDVQLQLEYLDMVNVNTRLF 367
QY 400 PPSPYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDA 459
DB 368 PVAMRLERVCCKDVE-----INGMFIPKGVVNMIPSVLHNDP-KYWTETP 411
QY 460 EEPFRWLEDVTSLSN-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRVRQV 518
DB 412 EKFLPERFSKKNKONIDPIYITPFG-----SGPRNCIGRFPALVNKKLALVRVLQNF 465

QY 519 EP 520
DB 466 KP 467

RESULT 12
A34101
cytochrome P450 3A5 - human
N;Alternate names: cytochrome P450 HLP2
N;Contains: oxidoreductase (EC 1.1.1.1)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: A34101; S06491; I52302
R;Aoyama, T.; Yamano, S.; Waxman, D.J.; Lapenson, D.P.; Meyer, U.A.; Fischer, V.; Tyndal
J. Biol. Chem. 264, 10388-10395, 1989
R;Schuetz, J.D.; Molowa, D.T.; Guzelian, P.S.
Arch. Biochem. Biophys. 274, 355-365, 1989
A;Title: Cytochrome P-450 hPCN3, a novel cytochrome P-450 IIA gene product that is diff
DNA-expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and cyclosporine.
A;Reference number: A34101; MUID:89278095; PMID:2732228
A;Accession: A34101
A;Molecule type: mRNA
A;Residues: 1-502 <AOY>
A;Cross-references: GB:J04813; NID:G181345; PIDN:AAA02993.1; PID:G181346
R;Schuetz, J.D.; Molowa, D.T.; Guzelian, P.S.
Arch. Biochem. Biophys. 274, 355-365, 1989
A;Title: Characterization of a cDNA encoding a new member of the glucocorticoid-responsi
A;Reference number: S06491; MUID:190025114; PMID:2802615
A;Accession: S06491
A;Status: not compared with conceptual translation

A;Molecule type: mRNA
A;Residues: 1-304, 'P', 306-317, 'F', 319-323, 'D', 325-376, 'G', 378-502 <SCH>
R;Jounaidi, Y.; Guzelian, P.S.; Maurel, P.; Villarem, M.J.
Biochem. Biophys. Res. Commun. 205, 1741-1747, 1994
A;Title: Sequence of the 5'-flanking region of CYP3A5: comparative analysis with CYP3A4
A;Reference number: I52302; MUID:95110318; PMID:7811260
A;Accession: I52302
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-24 <RES>
A;Cross-references: GB:S74699; NID:G786472; PIDN:AAD14157.1; PID:G4261857
C;Genetics:
A;Gene: GDB:CYP3A5
A;Cross-references: GDB:118783
A;Map position: 7q22.1-7q22.1
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monoxygenase;
F;302-463/Domain: cytochrome P450 homology <P45>
F;441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.5%; Score 385; DB 1; Length 502;
Best Local Similarity 28.1%; Pred. No. 1.5e-17;
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61
DB 7 LAVETWLLAVSLVLLYLGTRTHGLFKLGIQPTPLPFLGN---VLSYRQGLWKFDTE 63
QY 62 EHAHYREKYGSLRFPAGIAPVLSNDPKVFNHVM-KEAYDYPKPGMAARVLRATG-- 118
DB 64 CYKIRKVGWGT---YEG--QLPVLATDPVIRTVLVKECY-----SVFTNRSLSGPV 111
QY 119 ----DGVTVAAGBAHRRIRIMIPSLSAQAQVSMVPIFLEKGMELVDKMDAAEKDVA 174
DB 112 GFMKSAISLAEDBEWKIRSLSPFTSGKLKEMVPIIAQYGDVLRNLRRE-ASKG-- 167
QY 175 GESAGEKKATRLTETEGVDVKDWVGRATLDVNALAGFDYKSDSLN-----KTNLSLYVAF 228
DB 168 -----KPTVLKDIIFGAYSDMTITSTSGVSDSLNPNQDPPFVSTKFLKF 213
QY 229 VGLTGDGAPTLDSFKAIMWDFVFPYFRTMKRRHEIPLTQGLAVS---RRVGIELMEQKKQAV 286
DB 214 -----GFLDPL-FLSILPPFL-----TPVEALNVSLPFPKOTINFLSKS---- 252
QY 287 LGSASDAQVDKQVQGRDILSLVRANIAANLPESQKLSDEVLAAQISNLLFAGYETS 346
DB 253 VNRMKSRINDKQKRLDFLQMLDSQNSKQSTHKALESLELAAQSIIFIFAGYETS 312
QY 347 VLTWFMHRLSEDKAVQDKLREICQI---DTDMPTLDELNALPYLEAFVKESLRIDPPSPY 404
DB 313 VLSFTLYELATHPDVQVQVKEIDAVLPNKAPPTVDVQVQVYELDMVNVNTRLFPV 372
QY 405 ANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDA 464
DB 373 LERTCKKDVE-----INGVFIKGSMMVVIPTVALHNDP-KYWTETP 416
QY 465 ERLWLEDVTSLSN-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRVRQV 520
DB 417 ERFSKK-KDSIDPIYITPFG-----TGRNCIGRFPALVNKKLALIRVLQNFSP 466

RESULT 13
A29487
cytochrome P450 3A6 (version 1) - rabbit
N;Alternate names: cytochrome P450 3C
N;Contains: oxidoreductase (EC 1.1.1.1)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text_change 28-Jul-2000
C;Accession: A29487
R;Dalet, C.; Clair, P.; Daujat, M.; Fort, P.; Blanchard, J.M.; Maurel, P.
DNA 7, 39-46, 1988
A;Title: Complete sequence of cytochrome P450 3c cDNA and presence of two mRNA species w
A;Reference number: A29487; MUID:88166352; PMID:3349903

A;Accession: A29487
A;Molecule type: mRNA
A;Residues: 1-501 <DAL>
A;Cross-references: GB:M19139; NID:G165573; PIDN:AAA1430.1; PID:G165574
C;Genetics:
A;Gene: CYP3A6
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
F;301-462/Domain: cytochrome P450 homology <P45>
F;440/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.0%; Score 379; DB 2; Length 501;
Best Local Similarity 26.2%; Pred. No. 3.8e-17;
Matches 147; Conservative 91; Mismatches 208; Indels 114; Gaps 22;

Qy 19 ASTAFSLYLAPRSSLYNLOGNHTNY-FTGNFLDILSARTGEHAKYREKYGTSLRPA 77
Db 15 ASLVLLYLGXTSTHGLFKXGPIGPTPLPFIIGTILEYRKVINWDFIECRKXIG---KQW 71
Qy 78 GI--AGAPVLNSTDPKVFNVHM-KEAY-----DYPKPGMARVLRATGDGVVTAAGE 127
Db 72 GLFDGROPLAVITDPMIKTVLKECYSVFTNRRSFGPVGFMKKAVSI-----SEDE 123
Qy 128 AHKXRRIMTPISAQAVKSMVPIFEKGMELVDKMDAEDAKMDVAGESAGEKATRL 187
Db 124 DWKRVRLTSLPTTSGLKEMLPILAAQYGDVVLQNLROEAKSPST----- 170
Qy 188 TEGVDVVDKMGVGRATLDVMAAGFDYKSDSLQNTNE-----LYVAFVGLTDGFAPTL 239
Db 171 -----LKEIFGASMDVITGTSFGVWIDSLRNPQDPFVKVNRLLKPSF-----FDPLL 219
Qy 240 DSKALMDFVPYFTMKRRHRIPLTQGLVARRVGIEMEQKQAVLGSASDAQVKKD 299
Db 220 LSI-----TLFPFL-----TPIEALHIS-MFFKDVMDPLKTSVEKIKDKRKKQK 265
Qy 300 VQGRDILSLAVRANIANLPESCKLSDEVLQISNLPAGVETSTVLTWMEHRLSEDK 359
Db 266 -RRVDFIQLMINQSKREIDSHKALDIEVVAQSIILPAGYETTSITLSIFIMHLLATHP 324
Qy 360 AVQDKLREEICQIDTMP-----TLDELNALPYLEAFVKESLRDPPSPYANRECKDED 414
Db 325 DVQOKLQEE---IDTLNKLATYDTLVKMEYLDVWVNETLRLYPIAGRLERVCCKD 381
Qy 415 FIPLAEPIVGRDSVINEVRIKGTVMPLPLFNINRSKEIYGEDAEERPERNELEDVTS 474
Db 382 -----INGTFIPKGTIVM-PTVALHRDPOHWTE-PDEPRPERFSKQNDN 424
Qy 475 LNS-IEAPYGHQAFISGPRACFGWFAVAKKAFILVTLRRVQFE-----PT----- 521
Db 425 INPIYHPFG-----AGPENCILGMFALMNIKALVRLMQNFSFKLCKETQVPLKLGKQ 478
Qy 522 -ISHPEYEHITLISPRIV 540
Db 479 GLLOPEKPIVLKVVSSEGI 498

RESULT 14
A22631
Cytochrome P450 3A1, pregnenolone 16-alpha-carbonitrile-inducible - rat
N;Alternate names: testosterone 6beta-hydroxylase
N;Contains: unspecific monooxygenase (EC 1.14.14.1) cytochrome P450 PCN1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Aug-1987 #sequence revision 29-Aug-1987 #text change 28-Jul-2000
C;Accession: A22631; PX0035; S21697; S36137; S27107; S30378; I59218
R;Gonzalez, F.J.; Nebert, D.W.; Hardwick, J.P.; Kasper, C.B.
J. Biol. Chem. 260, 7435-7441, 1985
A;Title: Complete cDNA and protein sequence of a pregnenolone 16-alpha-carbonitrile-indu
A;Accession: A22631; MUID:85207783; PMID:3638989
A;Residues: 1-504 <GN>
A;Cross-references: GB:M10161; NID:G203777; PIDN:AAA1035.1; PID:G203778
R;Negata, K.; Gonzalez, F.J.; Yamazoe, Y.; Kato, R.

J. Biochem. 107, 718-725, 1990
A;Title: Purification and characterization of four catalytically active testosterone 6b
nally related forms. PX0032; MUID:90375438; PMID:2398038
A;Reference number: PX0035
A;Accession: PX0035
A;Molecule type: protein
A;Residues: 1-26 <NAG>
A;Experimental source: liver, Sprague-Dawley male rat
R;Lechner, M.C.
submitted to the EMBL Data Library, December 1991
A;Reference number: S21697
A;Accession: S21697
A;Molecule type: mRNA
A;Residues: 1-206, 'A', 208-212, 'I', 214-231, 'V', 233-504 <LEC>
A;Cross-references: EMBL:X64401; NID:G56038; PIDN:CAA45743.1; PID:G56039
R;Ribeiro, V.; Lechner, M.C.
Arch. Biochem. Biophys. 293, 147-152, 1992
A;Title: Cloning and characterization of a novel CYP3A1 allelic variant: Analysis of CY
A;Reference number: S36137; MUID:92117688; PMID:1731631
A;Accession: S36137
A;Molecule type: mRNA
A;Residues: 205-206, 'A', 208-212, 'I', 214-231, 'V', 233-234 <RIB>
A;Cross-references: EMBL:X64401
R;Telhada, M.B.; Pereira, T.M.; Lechner, M.C.
Arch. Biochem. Biophys. 298, 715-725, 1992
A;Title: Effect of dexamethasone and phenobarbital on run-on transcription rate and CYP
A;Reference number: S27107; MUID:93037516; PMID:1417000
A;Accession: S27107
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-24 <TEL>
A;Cross-references: EMBL:X62086
R;Cooper, K.O.; Reik, L.M.; Jayyosi, Z.; Bandiera, S.; Kelley, M.; Ryan, D.E.; Daniel,
Arch. Biochem. Biophys. 301, 345-354, 1993
A;Title: Regulation of two members of the steroid-inducible cytochrome P450 subfamily (
A;Reference number: S30378; MUID:93213168; PMID:7681660
A;Accession: S30378
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <COO>
R;Burger, H.
Proc. Natl. Acad. Sci. U.S.A. 89, 2145-2149, 1992
A;Title: Paradoxical transcriptional activation of rat liver cytochrome P-450 3A1 by de
to primary monolayer cultures of adult rat hepatocytes.
A;Reference number: I59218; MUID:92196074; PMID:1372436
A;Accession: I59218
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-24 <BUR>
A;Cross-references: GB:M86850; NID:G205919; PIDN:AAA41780.1; PID:G205920
C;Genetics:
A;Gene: CYP3A1; P450p
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
F;303-465/Domain: cytochrome P450 homology <P45>
F;443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.0%; Score 370; DB 2; Length 504;
Best Local Similarity 26.2%; Pred. No. 1.5e-16;
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

Qy 7 LTGALGLAFAFWASTAFSLY-LAPRRSSLYNLOG-----PHTNYFTGNF-LDI 54
Db 3 LLSALTLETWLAVLVLLVLLYGFGRTHGLFKKQIGPKPLPFGTVLVNYINGLWKFV 62
Qy 55 LSARTGEHAKYREKYGTSLRPAIGAPVLNSTDPKVFNVHM-KEAY-----DYPKP 106
Db 63 -----ECKKXGKING--LFDG--QNFPAITDTTEIKNVLYKECFSTNRRDFGPV 111
Qy 107 GMAARVLRATGDGVVTAAGEAHKRRIMTPISAQAVKSMVPIFEKGMELVDKMD 166
Db 112 GI-----MGKAVSVAKDEWKRYALLSPSTTSGLRKMFFPIIEQYGDILVKYLQ 163
Qy 167 AAEKDMVAGESAGEKATRLT-EGVDVVDKMGVGRATLDVMAAGFDYKSDSLQN----- 219

```

164 A-----ETGKPVTKKVGAGYSDVITSTSFVGVNVDLSLNNPKDPFV 204
220 -KTNEL-----YVAFVGLTDGFAPTLDSFKAINW--DFVPYFRTWKRRIEPL 264
205 EKTKKLLRFDPDFLFLSVLFPFLT-----PIYEMLNICMFPKXDSIEFFK- 250
265 TQGLAVSRVGIEMLEQKQAVLGSDQAVDKKQVQGRDILSLVRANIAANLPESQ-K 323
251 -----KVYRMKTRLDVSVQKRV-----DFLOMMNAENDSKDKESHTA 290
324 LSDBEVLAQISNLLFAGYETSSTVLTWMFRLSEDKAVQDKLREEICQI--DTDMPTLDE 381
291 LSDMEITAQSIIFIFAGVEPTSSLSFVLHSLATHPTQKQLQEEIDRALPNKAPTDT 350
382 LNALPYLEAFVKESLRLDPPSPYANRCLXDEDFIPLAEPVIGRDGSGVINEVRIKGTWV 441
351 VMEYELDMVLNETLRLYPIGNRLERVCKXKDE-----INGVFMPPKGSVV 395
442 MLPLFNINRSKFIVGEDAEERPERWLEDVTDLSNS-IEAPYGHQASFISGPRACFGMR 500
396 MIPSVLHSDQHWPE-PEERPERFSPKNGKSIDPVVLPFG-----NGPRNCIGMR 448
501 AVAEMKAFVTLRRVQPEP 520
449 ALMMNKLALTKVLQNFSPQ 468

RESULT 15
SS0892
Cytochrome P450 3A16 - mouse
N:Contains: oxidoreductase [EC 1.-.-.-]
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 09-May-1997 #text_change 28-Jul-2000
C:Accession: S50892
R:Itouh, S.; Sato, M.; Abe, Y.; Hashimoto, H.; Yanagimoto, T.; Kamataki, T.
Eur. J. Biochem. 226, 877-882, 1994
A:Title: A novel form of mouse cytochrome P450 3A (Cyp3a-16). Its cDNA cloning and expression
A:Reference number: S50892; MUID:95112853; PMID:7813478
A:Accession: S50892
A:Molecule type: mRNA
A:Residues: 1-504 <10>
A:Cross-references: EMBL:D26137; NID:G493670; PIDN:BA05133.1; PID:G666968
A:Experimental source: fetal liver
A:Note: In the authors' translation residues 9-20 do not match the nucleotide sequence
A:Note: the authors translated the codon TTC for residue 464 as Glu
C:Genetics:
A:Gene: CYP3a-16
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;
F:303-465/Domain: cytochrome P450 homology <p45>
F:469/Binding site: heme iron (Cys) (axial 1-legend) #status predicted

Query Match 12.9%; Score 368; DB 2; Length 504;
Best Local Similarity 25.8%; Pred. NO. 2e-16;
Matches 132; Conservative 100; Mismatches 171; Indels 108; Gaps 21;

QY 68 EXYGSTL-RFAGIAGAVLNSTDPKFNHVM-KEAY-----DYPKEGVAARVLRATG 118
DB 66 EXYKGTWGLFDG-QIFLFIITDPTIKNLVKECFSVFTNRQDFPFGIMGSISL--- 120
QY 119 DGWVTAEGEAKHRRIRIIPSLAQAVKSMVPIFLEKGMELVDKQMEDAAEKQWAGESA 178
DB 121 -----AKDEWKRYRALLSPFTSGNLKEMFPVIEQY-DILVKYLRQEAEGKQVPA--- 171
QY 179 GEKKATRLTEGVKDWVGRATLDYMALAGFDYKSDSLQKTNELVYAFVGLTDGFAPT 238
DB 172 -----VKDVLGAYSDVITSTFGVNIIDSLNPNP-----PF 203
QY 239 LDSFKAIM-WDF-----VPYFRTWKRRIEPLTQGLAVSRVGIEMLEQKQAVLGSA 290
DB 204 VENAKVLRVDFDPLSLVALFPFLTIYEM---LNICMFPKDSIEFFK----FYDRM 256

```

```

QY 291 SDQAVDKQVQGRDILSLVRANIAANLPESQ--LSDBEVLAQISNLLFAGYETSSTVLT 349
DB 257 TENRLDSKQKRVDFIYLMEAYNKSOKSHKALSEIEITAQSIFIFAGVEITSSILS 316
QY 350 WMFRLSEDKAVQDKLREEICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANR 407
DB 317 FTVYSLATHPDIOKQLQEEIDEALPNKAPTDTVMAMEYLDMLNETLRLYPITNRLQR 376
QY 408 ECLKDEDFIPLAEPVIGRDGSGVINEVRIKGTWVMLPLFNINRSKFIVGEDAEERPER 467
DB 377 VKKQVE-----INGYIFKGSVTIIPSVLHSDQHWPE-PEEFQPERF 420
QY 468 LEDVTDLSNS-IEAPYGHQASFISGPRACFGMRPAVEMKAFVTLRRVQPEPIISHPE 526
DB 421 SKENKGSIDPVVLPFG-----NGPRNCIGMRFALMMNKLALIKVLQNFSPQC---K 470
QY 527 YEHTIILISPRIVGREKEGYQMRLO-VKPV 556
DB 471 ETQIPLKLSR-----ELLQPVAPI 490

```

Search completed: April 2, 2004, 14:01:26
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:59:14 ; Search time 12 Seconds
(without alignments)
871.939 Million cell updates/sec

Title: US-10-066-007-1
Perfect score: 2852
Sequence: 1 MFVLVLLTGALGAASWAS.....RIVGREKEGYQNRLOVKPVE 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 155783 seqs, 18785033 residues

Total number of hits satisfying chosen parameters: 155783

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New: *
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pdp.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pdp.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pdp.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pdp.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pdp.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pdp.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	370	13.0	504	6	US-10-784-004-427		Sequence 427, App
2	370	13.0	504	6	US-10-784-004-948		Sequence 948, App
3	348.5	12.2	546	6	US-10-275-027A-405		Sequence 405, App
4	347.5	12.2	524	6	US-10-275-027A-877		Sequence 877, App
5	347.5	12.2	524	6	US-10-275-027A-878		Sequence 878, App
6	347.5	12.2	524	6	US-10-275-027A-879		Sequence 879, App
7	345.5	12.1	522	6	US-10-767-701-46006		Sequence 46006, A
8	336	11.8	539	6	US-10-767-701-45390		Sequence 45390, A
9	328	11.5	466	1	PCT-US04-04280-33		Sequence 33, Appl
10	323.5	11.3	568	6	US-10-275-027A-406		Sequence 406, App
11	310.5	10.9	470	6	US-10-776-871-12		Sequence 12, Appl
12	310.5	10.9	496	6	US-10-776-871-10		Sequence 10, Appl
13	310.5	10.9	496	6	US-10-782-695-19		Sequence 19, Appl
14	305.5	10.7	253	6	US-10-767-701-40524		Sequence 40524, A
15	286	10.0	276	6	US-10-767-701-41801		Sequence 41801, A
16	282.5	9.9	374	6	US-10-767-701-45728		Sequence 45728, A
17	281.5	9.9	505	6	US-10-776-871-2		Sequence 2, Appl
18	277.5	9.7	429	6	US-10-275-027A-407		Sequence 407, App
19	272	9.5	368	6	US-10-767-701-45321		Sequence 45321, A
20	257.5	9.0	509	6	US-10-100-683-8244		Sequence 8244, Ap
21	257.5	9.0	509	6	US-10-152-372-108		Sequence 108, App
22	243.5	8.5	216	6	US-10-767-701-41930		Sequence 41930, A
23	240.5	8.4	209	6	US-10-767-701-39467		Sequence 39467, A
24	238	8.3	413	6	US-10-767-701-43099		Sequence 43099, A
25	228.5	8.0	537	6	US-10-782-695-17		Sequence 17, Appl
26	217.5	7.6	195	6	US-10-767-701-44683		Sequence 44683, A

Sequence 455, App
Sequence 39328, A
Sequence 46524, A
Sequence 45702, A
Sequence 40358, A
Sequence 659, App
Sequence 880, App
Sequence 8, Appl
Sequence 43383, A
Sequence 5, Appl
Sequence 8400, Ap
Sequence 38337, A
Sequence 39135, A
Sequence 46653, A
Sequence 879, App
Sequence 42166, A
Sequence 100, App
Sequence 47069, A
Sequence 456, App

ALIGNMENTS

RESULT 1

US-10-784-004-427

; Sequence 427, Application US/10784004

; GENERAL INFORMATION:

; APPLICANT: Biogen Idec

; TITLE OF INVENTION: Surrogate Markers of Pain

; FILE REFERENCE: 08201.6029-00000

; CURRENT APPLICATION NUMBER: US/10784,004

; CURRENT FILING DATE: 2004-02-20

; NUMBER OF SEQ ID NOS: 1251

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 427

; LENGTH: 504

; TYPE: PRT

; ORGANISM: rat

US-10-784-004-427

Query Match 13.0%; Score 370; DB 6; Length 504;

Best Local Similarity 26.2%; Pred. No. 1.1e-20;

Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY	7	LTGALGAASWASIAFFSLY-LAPRRSLYNLQ-----PNTNYTGNF-LDI 54
Db	3	LLSALTLETWLLAVLVLLYGFTRHGLFKQGPGLPFFGTVLYNYGLWKFDV 62
QY	55	LSARTGEBHAKYREKYGSTLRFPAGIAGAPVLNSTDPKFNHYM-KEAY-----DYKPK 106
Db	63	-----ECHKKYKING---LFDG--QMPLFAITDTEMIKNVLKCEFSVFTNRDRFGPV 111
QY	107	GMAARVLRIATGCGVTVTAEGEAHRRHRIIMPSLSAQAVKSWVPFLEKGMELVDKMD 166
Db	112	GI-----MGRKAVSVAKDEEWKRYALLSPFTTSGRLEKMFPIIEQYGDILVKYLOE 163
QY	157	AAEKDMAVGESAGEKAKTRLET-EGVDVKDWVGRATLDVVALAGFDYKSDSLQN----- 219
Db	164	A-----ETGKPTMKKVFAGYSDMTITSTSGVGVVDSLNNPKDPFV 204
QY	220	-KTNEL-----YVAFVGLTGDGFAPTLDSFKAIMW--DFVPYFRTMKRHEIFL 264
Db	205	EKTKLLRFDFDPLFLSVLFPFLT----PIVEMLNICMFPKDSTIEFFK----- 250
QY	265	TQGLAVSRVUCIELMEQKQAVLGSASDAQVDKQVGGDRILSLVRAIANALPESQ-K 323
Db	251	-----KPYMKKTRLDSVQKRV-----DFLQMMNANDSKDKESHYA 290
QY	324	LSDEVLIAQISNLLFAGYETSSVLTWFMFRLSEKAVQDKLREIICQI--DTDMPTLDE 381
Db	291	LSDEMIQAQSIIFIFAGYETSSVLSVLHSLATHPTDKLQEEIDRALPNKAPTDTOT 350

RESULT 4

US-10-275-027A-877
; Sequence 877, Application US/10275027A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP3/US
; CURRENT APPLICATION NUMBER: US/10/275,027A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 877
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-275-027A-877

Query Match 12.2%; Score 347.5; DB 6; Length 524;
Best Local Similarity 23.6%; Pred. No. 6.1e-19;
Matches 140; Conservative 104; Mismatches 211; Indels 137; Gaps 26;
QY 3 ILVLTGALGLA-AFSWASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDILSARTG- 60
Db 21 LLLLVLGWSLLARILAWT---YAFYNNCR--LQCFPPQPKRNWFWGLHGLITPEEGL 74
QY 61 ----BEHAKYREKYGSTLRPAGIAGAPVNSTDPKVFNVHMK-EAYDYKPGMAARVLEI 115
Db 75 KSTQMSATYSQ--GFTVWLGPI--IPFIVLCHPDITRSITNASAAIAPKDNLFRFLKP 130
QY 116 ATGDGVVTAEGEAHRRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAG 175
Db 131 WLGEGLLLSGGDKWSRHRRLTAPAFHNILKSYITFNKSNIMLDKWOHLASE-----G 185
QY 176 ESAGEKKAIRLETGVVDKDWVGRATLDVWALAGFDYKSDSLQNKTNELVYAFVGLTDF 235
Db 186 SSC-----LDMFEHISLMTLDSLQKCIFFSDS-HCOERPSE-YIA----- 223
QY 236 APTLDSFKAIM-----WDFVPYFR-----TMKRRHEIPLT 265
Db 224 --TILELSALVEKRSQHILQHMDFLYLSDHGRFRHACRLVHDFTDVIRERRRTP-T 280
QY 266 QGLAVSRVGVTELMCKEQAVLGSASQAVDKVDQVGRDILSLVLRANLANLPESOKLS 325
Db 281 QGI-----DFFKDKAKSKTLDFIDVL-----LISKDEGKALS 314
QY 326 DEVLIAQISNLLFAGYETSTVLTWMPHRLSEDAVQDKLREBICQI--DTDMPTL--DE 381
Db 315 DEDIRAEADTFMFGGHDITAGSLWSVLYNLARHPEYQRCRQEVQVQELLKDRDPKEIEMDD 374
QY 382 LNALPYLEAFVKESLRIDPPSPYANRECLKDEDFIPLAEFVIGRDSGVINEVRIKGTWV 441
Db 375 LAQLPFLTWCKESLURLHPPAFFISRCTQD-----IVLPDGRV-----IPKGITC 420
QY 442 MLPLFNINRSFIYGEDAEERPERWLEVDVDSLNSI-EAPYGHQASFISSGPRACFGWRF 500
Db 421 LLDIIGVHNPTVM-PDEVYDPERF-----DPENSKGRSLAP-IPFSAGPRNCIGQAF 473
QY 501 AVAEMKALFVTLRRVQEPILSHPEVHITLIISRPVIGREKEGYQMLQ 552
Db 474 AAEMKVVIALMLLHFRFLPDHTEPR-----RKLELINRAEGGLWURVE 517

RESULT 5

US-10-275-027A-878
; Sequence 878, Application US/10275027A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP3/US
; CURRENT APPLICATION NUMBER: US/10/275,027A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 878
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-275-027A-878

Query Match 12.2%; Score 347.5; DB 6; Length 524;
Best Local Similarity 23.6%; Pred. No. 6.1e-19;
Matches 140; Conservative 104; Mismatches 211; Indels 137; Gaps 26;
QY 3 ILVLTGALGLA-AFSWASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDILSARTG- 60
Db 21 LLLLVLGWSLLARILAWT---YAFYNNCR--LQCFPPQPKRNWFWGLHGLITPEEGL 74
QY 61 ----BEHAKYREKYGSTLRPAGIAGAPVNSTDPKVFNVHMK-EAYDYKPGMAARVLEI 115
Db 75 KSTQMSATYSQ--GFTVWLGPI--IPFIVLCHPDITRSITNASAAIAPKDNLFRFLKP 130
QY 116 ATGDGVVTAEGEAHRRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAG 175
Db 131 WLGEGLLLSGGDKWSRHRRLTAPAFHNILKSYITFNKSNIMLDKWOHLASE-----G 185
QY 176 ESAGEKKAIRLETGVVDKDWVGRATLDVWALAGFDYKSDSLQNKTNELVYAFVGLTDF 235
Db 186 SSC-----LDMFEHISLMTLDSLQKCIFFSDS-HCOERPSE-YIA----- 223
QY 236 APTLDSFKAIM-----WDFVPYFR-----TMKRRHEIPLT 265
Db 224 --TILELSALVEKRSQHILQHMDFLYLSDHGRFRHACRLVHDFTDVIRERRRTP-T 280
QY 266 QGLAVSRVGVTELMCKEQAVLGSASQAVDKVDQVGRDILSLVLRANLANLPESOKLS 325
Db 281 QGI-----DFFKDKAKSKTLDFIDVL-----LISKDEGKALS 314
QY 326 DEVLIAQISNLLFAGYETSTVLTWMPHRLSEDAVQDKLREBICQI--DTDMPTL--DE 381
Db 315 DEDIRAEADTFMFGGHDITAGSLWSVLYNLARHPEYQRCRQEVQVQELLKDRDPKEIEMDD 374
QY 382 LNALPYLEAFVKESLRIDPPSPYANRECLKDEDFIPLAEFVIGRDSGVINEVRIKGTWV 441
Db 375 LAQLPFLTWCKESLURLHPPAFFISRCTQD-----IVLPDGRV-----IPKGITC 420
QY 442 MLPLFNINRSFIYGEDAEERPERWLEVDVDSLNSI-EAPYGHQASFISSGPRACFGWRF 500
Db 421 LLDIIGVHNPTVM-PDEVYDPERF-----DPENSKGRSLAP-IPFSAGPRNCIGQAF 473
QY 501 AVAEMKALFVTLRRVQEPILSHPEVHITLIISRPVIGREKEGYQMLQ 552
Db 474 AAEMKVVIALMLLHFRFLPDHTEPR-----RKLELINRAEGGLWURVE 517

RESULT 6

US-10-275-027A-879
 ; Sequence 879, Application US/10275027A
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tang et al
 ; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 785CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/275,027A
 ; CURRENT FILING DATE: 2002-10-30
 ; PRIORITY APPLICATION NUMBER: PCT/US01/02623
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 09/491,404
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: 09/617,746
 ; PRIOR FILING DATE: 2000-07-17
 ; PRIOR APPLICATION NUMBER: 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: 09/633,870
 ; PRIOR FILING DATE: 2000-09-15
 ; NUMBER OF SEQ ID NOS: 944
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 879
 ; LENGTH: 524
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-275-027A-879

Query Match 12.2%; Score 347.5; DB 6; Length 524;
 Best Local Similarity 23.6%; Pred. No. 6, 1e-19;
 Matches 140; Conservative 104; Mismatches 21; Indels 137; Gaps 26;
 QY 3 ILVLTGALGLA-AFSWASIAFSLYLAAPRSSLYNQGNHNYFTGNFLDILSARTG- 60
 Db LULLVVGSLARILAWT-----YAFYNNCR--LQCFPPPKRNWFGHLGLITPTEGL 74
 QY 61 ----EEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVENHMK-EAYDYPKEGMAARVLR 115.
 Db 75 KDSQMSATYSQ--GFTVNLGPI--IPFIVLCHDPTTRISITNAAAPKDNIFIRFLKP 130
 QY 116 ATGDGVVTAAGBAHKKRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKXMBDAEKDMVAG 175
 Db 131 WLGEIGILLSGDGKWSRRHRLTPAFHNLKSYITFNKSNANIMLDKQWHLASE-----G 185
 QY 176 ESAGEKKAATLRTGEGVDYKDWVGRATLDVMALAGFDYKSDSLQNKTNELVAVGLTDGF 235
 Db 186 SSC-----LDNFHISLMTLDSLOKCFISFDS-HCQERPSE-YIA----- 223
 QY 236 APTLDSFKAIM-----WDFVYFPR-----TMKRRHEIPT 265
 Db 224 --TILELSALVEKRSQHILQHMDFLYLSHDGRFRHACRLVHDFTDAVIRERRTLPT 280
 QY 266 QGLAVSRVVGIELMEQKQAVLGASDAQVDKQDVQGRDILSLVRANIANIAPESQKLS 325
 Db 281 QOI-----DFFKDKAKSKTLDLFDVL-----LLSKDEDGKALS 314
 QY 326 DEVLIAQISNLLFAGYETSSVLTWFMHRLSEDAVQDKLREIEICQI--DTDMPTL--DE 381
 Db 315 DEDIRAEADTFMGGDDHTASGLSWLYNLARHPEYQECRQEQVELLKDRDPKEIEND 374
 QY 382 LNALPYLEAFVKESLRDPPSPYANRECLDEDFIPLAEVPIGRDGSVINEVRIKGTWV 441
 Db 375 LAQPFLLTMCVKESLRDPPSPYANRECLDEDFIPLAEVPIGRDGSVINEVRIKGTWV 441
 QY 442 MLPLFNINRSKFYIGEDAEERPRERWLEDVTDLSLSNI-EAPYGHQASFTSGPRACFGWRF 500
 Db 421 LIDIIGVHNHTVW-POPEVYDPRF-----DENSKGRSPLAF-IPFAGAPNCGIQAF 473
 QY 501 AVAEKARLFTVLRVQPEPIIISHPHEVHTLISRPVIGREKEGYQMRLQ 552
 Db 474 AMAEKVVVALMLLHFRFLPDHTPR-----RKLEIMRAEGGLWLRVE 517

RESULT 7

US-10-767-701-46006
 ; Sequence 46006, Application US/10767701
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-23
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 46006
 ; LENGTH: 522
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2930_1.pap
 ; US-10-767-701-46006

Query Match 12.1%; Score 345.5; DB 6; Length 522;
 Best Local Similarity 22.9%; Pred. No. 8, 6e-19;
 Matches 134; Conservative 90; Mismatches 21; Indels 149; Gaps 21;
 QY 6 LLTGALGLAAPSASIA-AFFSLVLAPEP-SSLYNQGNHNYFTGNFLDILSARTGEEH 63
 Db 18 LLOGLLAM-LIANGSYOAAETFWLRPRRLNRLRAQGLSGTEY-----LPFAGDLKEN 69
 QY 64 AKYREKYGSTLRFAGIAGAPVLNSTDPKVENHVMKEAYDY-----PKFGM-----AAR 111
 Db 70 DRLANDE-----ARSRMPLSHDIVPRVMPHLFNTVKEHGNISITWFGPIPRVITTEAE 122
 QY 112 VLRLATGD-----GVTAAEGSAHKKRRIRIMPSLSAQAVKSMVPI 151
 Db 123 LIRDVLNKGSHFEPKFSINRKLKIMVIALGASHDGEKAKWRHRLNPAFLHKLKMLFA 182
 QY 152 FLEKGMELVDKMBDAEKDMVAGESAGEKKAATRLTEGVVDKDWVGRATLDVMALAGFD 211
 Db 183 PSTCSTELIDRWENKLAASD-----GSVELDIC----- 210
 QY 212 YKSDSLQNKTNELVAVGLTDGFATLDSFKAIMWDFVYFRTMKRRHEIPTOGGLAV- 270
 Db 211 -----PPELHPT-----PGMIPPPSRGSSLLG-----RTLCLRRGQAGALIAL 252
 QY 271 -----SRRVGIELMEQKQAVLG--SASDAQVDKQDVQGRDILSLVRANI 314
 Db 253 HLFPFPGGFFPPPHNRT-----RGRKRRLRGMIKREBAIENGSTSGNDLLGLLQSNM 307
 QY 315 AANLPESQKLSDEVLIAQISNLLFAGYETSSVLTWFMHRLSEDAVQDKLREIEICQI-- 372
 Db 308 DSG-KGSLRMTSEDEVECKLFYFAGMETTSVLLTWTVLGMHPPEWODRAREEVLVFG 366
 QY 373 DTDMPITDELNALPYLEAFVKESLRDPPSPYANRECLDEDFIPLAEVPIGRDGSVINE 432
 Db 367 RDKQPNFDGJGRLLKVTMILYEVRLYPPAVSLNRTFKD-----MTIGG 411
 QY 433 VRTKGMWMLPLFNINRSKFYIGEDAEERPRERWLEDVTDLSLSIEA--PYGHQASFTS 490
 Db 412 ISVPAGVILELPTIVVHHSTDYMGKDAHEFKERFAEGISKATKORPAFFPFGW----- 465
 QY 491 GPRACFGWRFAVAKMAFLVTLRRVQPE--PIISHPHEVHTL 532
 Db 466 GPRICIGQNFALLEAKWALSMLQRFOLSPSYTHAPYTVLTL 509

RESULT 8

US-10-767-701-45390
 ; Sequence 45390, Application US/10767701
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.

[illegible]

QY 453 RTYGDAEERPERWLEDVTDLSNIRAPYGHQASPTSGPRACGHRFAVAEMKAFLEVT 512
 Db 378 AWW-PDPEVYDFRF-----DPKIKERSPLAFIPFSGAPRNCIGQAFAMKVVGLT 431
 QY 513 LRVQFEPIIHPVEYEHITLIISRPVIGREKEGYQMRLO 552
 Db 432 LURFRALPDHTEPR-----RKPELVRAEGGLMLRVE 463

RESULT 10
 US-10-275-027A-406
 ; Sequence 406, Application US/10275027A
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tang et al.
 ; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 785CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/275,027A
 ; PRIOR FILING DATE: 2002-10-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/02623
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 09/491,404
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: 09/517,745
 ; PRIOR FILING DATE: 2000-07-17
 ; PRIOR APPLICATION NUMBER: 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: 09/633,870
 ; PRIOR FILING DATE: 2000-09-15
 ; NUMBER OF SEQ ID NOS: 944
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 406
 ; LENGTH: 568
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-275-027A-406

Query Match 11.3%; Score 323.5; DB 6; Length 568;
 Best Local Similarity 22.2%; Pred No. 4.6e-17;
 Matches 137; Conservative 105; Mismatches 210; Indels 165; Gaps 26;

QY 3 ILVLLTGAALGA-AFSWASIAFSLYLAPRRSSLYNLOGPNHTYFTGNFDILSARTG- 60
 Db 43 LLLLVGSSNLLARILAWT---YAFVNNCR--LQCFPPKPRNFWGHLIPTBEGL 96

QY 61 -----EEHAKYREKYGSTLRFAGIAGAPVLNSDTPKVFNVK- EAYDYKPGMAARVLR 115
 Db 97 KNTQMSATYSQ--GFTIWLGEI--IPFVLCHPDTIRSIITNAGAAIAPKDNLFIRELKP 152

QY 116 ATGDGVVTAEGEAHKEHRRIMIPSLSAQAQVSMVPIFLEKGMELVDKMMEDAAEKDMAVG 175
 Db 153 WLCEGILLSGGDKWSRRHMLTAPAFHNILKSYITFNKSNIMLDKQWHLASE-----G 207

QY 176 ESAGEKATRETEGVVDKVDWVGRATLDVMAAGFDYKSDLSQNTNELYVAFVGLTDGF 235
 Db 208 SSC-----LDMFEHLSMTLDSLOKCIFFSOS-HCQEPSE-YIA----- 245

QY 236 APTLDSFKAIM-----WDFEYFER-----TMKRREHIEPLT 265
 Db 246 --FILELSALVEKRSQHILQHMDFLYLSHDGRFRHACRLVHDFDAVIERRRTLP-T 302

QY 266 QGLAVSRVGIEMEQKQAVLGASDAQVDKQVQGRDILSLVRANIANLSPESQKLS 325
 Db 303 QGI-----DDFFKDKAKSKTLDFIDVL-----LLSKDSDGKALS 336

QY 326 DEEVLQISNLLF-----AGVETSTVLTMFHRLSLSDKAVQD 363
 Db 337 DEIDRAEDTFMGGPQYLGAHVPPVLPKSLPGCSSGHDITASGLSWLYNLARHPPEQE 396

QY 364 KIREICQI--DTDMPTL--DEINALPYLEAFVKLSRLDPPSPYANRECUKDEDFIPLA 419
 Db 397 RCRQEVQELKDRPKPEIENWDDLAQPLFLTMCVKESLSRLHPPAPFISRCCTQD----- 449

QY 420 EPVICRDSGVINEVRITKGTMTVMPLFNINRSKFIYGE-----DAEEFRPERWLEDVTDLSL 475
 Db 450 --IVLPDGRV-----IPKGITCLIDIIVGHNPTVMPDGVYDPRFDPDEN----- 493

QY 476 NSIEAPYGHQASPTSGPRACGHRFAVAEMKAFLEVTLSRVQFEPIIHPVEYEHITLIIS 535
 Db 494 SKGRSLPAP-IPFSGAPRNCIGQAFAMKVVGLMLLHLFRFLPDHTEPR-----R 544

QY 536 RPRIVGREKEGYQMRLO 552
 Db 545 KLEILMRAEGGLMLRVE 561

RESULT 11
 US-10-776-871-12
 ; Sequence 12, Application US/10776871
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria Alexandra
 ; APPLICANT: Meyers, Rachel
 ; APPLICANT: Rudolph-Owen, Laura A.
 ; TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF
 ; FILE REFERENCE: 10448-193001
 ; CURRENT APPLICATION NUMBER: US/10/776,871
 ; CURRENT FILING DATE: 2004-02-11
 ; PRIOR APPLICATION NUMBER: US/10/175,896
 ; PRIOR FILING DATE: 2002-06-20
 ; PRIOR APPLICATION NUMBER: 10/067,668
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: 60/286,140
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 09/823,901
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/10720
 ; PRIOR FILING DATE: 2001-04-02
 ; PRIOR APPLICATION NUMBER: 60/193,920
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/862,658
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: PCT/US01/16380
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/205,675
 ; PRIOR FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/882,837
 ; PRIOR FILING DATE: 2001-06-15
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 470
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: consensus sequence
 US-10-776-871-12

Query Match 10.9%; Score 310.5; DB 6; Length 470;
 Best Local Similarity 24.7%; Pred No. 3.5e-16;
 Matches 141; Conservative 103; Mismatches 172; Indels 155; Gaps 28;

QY 24 FSLYLAPRRSSLYNLOGPNHTYFTGNFDILSARTGEEHAKYREKYGSTLRFAGIAGAP 83
 Db 15 FTYLGLPR--PWVLTGPEAVK-----EVLIDK-GEFPAKGRGDFNPTF-----P 56

QY 84 VLNSDTPKVFNVHVMKEAYDYKPGVAAARVLRATGDDGVVTAEGEAHKEHRRIMIPSL--- 140
 Db 57 WLUSK-----GYREQGL-----LFDNGPKWKRLARFSLTLRPH 90

QY 141 -----SAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMAVGESAGEKATRLTEGVDK 194
 Db 91 FCGAVYSKESQKLEFPRIQEEARDLVERLRK-----EQAGSP-----IDIT 131

QY 195 DWVGR-ATLDVMAAGFDYKSDLSQNK-----TNELYVAFVGLTDGFAP---TL 239

Db 132 ELLARLAPLNVICSLFGVRFYDLRPEDEPEFLKLDKLLNEMF-----DRVSPWHQLL 184
 Qy 240 DSPKALMWDVFP---VFRTMKRRHEIPLTQGLAVSRVGIEMEOKQAVLGSDQAVD 296
 Db 185 DIPFLL-RYLPGLSLFRKAFKAKDL-----KDYLDKLIBERETL-----EPAGD 229
 Qy 297 KDVQGRDILSLVRANIANLPESOKLSDEEVLQIISNLLFAGVETSTVLTWTFHRLS 356
 Db 230 PRRLDITGFLDSLLEAKREGNPKSE-LSDEELAATVLDLLFAGTETTSSTLSWALYLLA 288
 Qy 357 EDKAVQDKLREETICQI--DTDMPT--LDELNALPYLEAFVKESLKDPPSPYANRECLD 412
 Db 289 KHPEVQAKLREEIDEVIGRDSRSTYDVARAQMPYLDVAVIKETLRLYP----- 336
 Qy 413 EDRTPLAEP-VIGRDSGVINEVRITKGTWMLPLFNINRSKIYGEDAEPEPRMLEDV 471
 Db 337 --VVPDLLPRVATKOTEIPDGYLIPKGTIVLVNLSLHRDPKVF-PNPEEDPDPFLDEN 393
 Qy 472 TDSLNSIE-APYGHQASFSIGRACFGWRFAVAKMKAFLVTILRRVQPEPIISHPEYE-- 528
 Db 394 GKPKKSYAFLPFG-----AGPRNCLGERLARMELFLATLILQR-----PPELELA 439
 Qy 529 ----HITLIISPRIVGREKEGYQMLQVQP 555
 Db 440 VPPGDIPSLTPKPEL-GLPSKPLYKVQLRP 469

RESULT 12
 US-10-776-871-10
 ; Sequence 10, Application US/10776871
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria A.
 ; APPLICANT: Meyers, Rachel
 ; APPLICANT: Rudolph-Owen, Laura A.
 ; TITLE OF INVENTION: NOVEL HUMAN ENZYME FAMILY MEMBERS AND USES THEREOF
 ; FILE REFERENCE: 10448-193001
 ; CURRENT APPLICATION NUMBER: US/10/776,871
 ; CURRENT FILING DATE: 2004-02-11
 ; PRIOR APPLICATION NUMBER: US/10/175,696
 ; PRIOR FILING DATE: 2002-06-20
 ; PRIOR APPLICATION NUMBER: 10/067,668
 ; PRIOR FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: 60/266,140
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 09/823,901
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/10720
 ; PRIOR FILING DATE: 2001-04-02
 ; PRIOR APPLICATION NUMBER: 60/193,920
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/862,658
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: PCT/US01/16380
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/205,675
 ; PRIOR FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/882,837
 ; PRIOR FILING DATE: 2001-06-15
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 496
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: consensus sequence
 US-10-776-871-10
 Query Match 10.9%; Score 310.5; DB 6; Length 496;
 Best Local Similarity 24.7%; Pred. No. 3.8e-16;
 Matches 141; Conservative 103; Mismatches 172; Indels 155; Gaps 28;

Qy 24 FSLYLAPRRSLYNLQGNHTNYFTGNFELDILSARTGEEHAKYREKYGSLTRFAGIAGAP 83
 Db 41 FLYLIGPR--PVVLTJGEAVK-----EVLIDK-GEEFAKGRGDNPTF-----P 82
 Qy 84 VLNSTDTPKVFNVHMKAEADYKPKGMAARVLRATGDDGVVTAEGEAHKHRRIMTIPSL--- 140
 Db 83 WLSK-----GYREQGL-----LFSNKGPKWKRLRRFSLTLTTRFH 116
 Qy 141 ----SQAQVSMVPIFLEKGMELVDKMDAAEKMAVGESAGEKATRLTEGVDDVK 194
 Db 117 FGWGYSKESQKLEPRIOEABDLVERLRK-----EQAGSP-----IDIT 157
 Qy 195 DWVGR-ATLDVVALAGFDYKSDSLQK-----TNELYVAFVGLTQGFAP---TL 239
 Db 158 ELLARLAPLNVICSLFGVRFYDLRPEDEPEFLKLDKLLNEMF-----DRVSPWHQLL 210
 Qy 240 DSPKALMWDVFP---VFRTMKRRHEIPLTQGLAVSRVGIEMEOKQAVLGSDQAVD 296
 Db 211 DIPFLL-RYLPGLSLFRKAFKAKDL-----KDYLDKLIBERETL-----EPAGD 255
 Qy 297 KDVQGRDILSLVRANIANLPESOKLSDEEVLQIISNLLFAGVETSTVLTWTFHRLS 356
 Db 256 PRRLDITGFLDSLLEAKREGNPKSE-LSDEELAATVLDLLFAGTETTSSTLSWALYLLA 314
 Qy 357 EDKAVQDKLREETICQI--DTDMPT--LDELNALPYLEAFVKESLKDPPSPYANRECLD 412
 Db 315 KHPEVQAKLREEIDEVIGRDSRSTYDVARAQMPYLDVAVIKETLRLYP----- 362
 Qy 413 EDRTPLAEP-VIGRDSGVINEVRITKGTWMLPLFNINRSKIYGEDAEPEPRMLEDV 471
 Db 363 --VVPDLLPRVATKOTEIPDGYLIPKGTIVLVNLSLHRDPKVF-PNPEEDPDPFLDEN 419
 Qy 472 TDSLNSIE-APYGHQASFSIGRACFGWRFAVAKMKAFLVTILRRVQPEPIISHPEYE-- 528
 Db 420 GKPKKSYAFLPFG-----AGPRNCLGERLARMELFLATLILQR-----PPELELA 465
 Qy 529 ----HITLIISPRIVGREKEGYQMLQVQP 555
 Db 466 VPPGDIPSLTPKPEL-GLPSKPLYKVQLRP 495

RESULT 13
 US-10-782-695-19
 ; Sequence 19, Application US/10782695
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria A.
 ; APPLICANT: Curtis, Rory A.J.
 ; APPLICANT: Tsai, Fong-Ying
 ; APPLICANT: Hodge, Martin R.
 ; APPLICANT: Meyers, Rachel E.
 ; APPLICANT: MacBeth, Kyle J.
 ; APPLICANT: Bandaru, Rajasekhar
 ; TITLE OF INVENTION: NOVEL 14275, 54420, 8797, 27439, 68730.
 ; TITLE OF INVENTION: 69112 AND 52908 MOLECULES AND USES THEREFOR
 ; FILE REFERENCE: MP104-0020NMIM
 ; CURRENT APPLICATION NUMBER: US/10/782,695
 ; CURRENT FILING DATE: 2004-02-19
 ; PRIOR APPLICATION NUMBER: US 09/945,254
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: US 60/229,829
 ; PRIOR FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: US 09/945,301
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: US 60/229,301
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: US 10/007,399
 ; PRIOR FILING DATE: 2001-11-05
 ; PRIOR APPLICATION NUMBER: US 09/390,039
 ; PRIOR FILING DATE: 1999-09-03
 ; PRIOR APPLICATION NUMBER: US 09/146,416
 ; PRIOR FILING DATE: 1998-09-03
 ; PRIOR APPLICATION NUMBER: US 10/024,036
 ; PRIOR FILING DATE: 2001-12-17

QY	332	QISNLLFAGYEYTSVLTWFMHRLSEDKAVODKLRBEICOI--DTDMPITDELNALPYLEA	390
D5	79	ECKLFYFAGMETTVLLTWTMLLSMHPEWDAREEVLALFGKNQGYDGLSRLLKVTVM	138
QY	391	FVKESURLDPSPYANRECLDKEDFIPLAEPIVGROGSVINEVRITGTMWMLPFINNR	450
D5	139	ILYEVLRLYPPAIAPASRKYTE-----MWIGDYTPAGVIVELPWFVIHH	183
QY	451	SKPEICYDEABERPERRWLEDYTDLSINIEA--PVGHQAISGPACFCGWREFAVAEMKAF	508
D5	184	DADIGSDVHEFRPERFAEGIAKSORLAFPFPGW-----GPRICIQNTALLAEAKWA	237

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 14:00:49 ; Search time 40 Seconds
(without alignments)
3653.561 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFTLVLLTGALGLAFAFWSAS.....RIVGREKGYQMLQKPFVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1070241 seqs, 262374223 residues

Total number of hits satisfying chosen parameters: 1070241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2852	100.0	557	US-10-066-007-1	Sequence 1, Appli
2	2852	100.0	557	US-10-066-007-3	Sequence 3, Appli
3	398.5	14.0	503	US-10-313-963A-56	Sequence 56, Appl
4	396.5	13.9	503	US-10-146-575-2	Sequence 2, Appli
5	385.5	13.5	503	US-09-957-997-3	Sequence 3, Appli
6	384	13.5	537	US-10-425-114-38180	Sequence 38180, A
7	377.5	13.2	527	US-10-425-114-67055	Sequence 67055, A
8	376.5	13.2	547	US-10-425-114-65616	Sequence 65616, A
9	368.5	12.9	547	US-10-425-114-61218	Sequence 61218, A
10	368.5	12.9	560	US-10-425-114-59349	Sequence 59349, A
11	368.5	12.9	562	US-10-425-114-59350	Sequence 59350, A
12	367	12.8	520	US-09-992-901-2	Sequence 2, Appli
13	365.5	12.8	662	US-10-425-114-38827	Sequence 38827, A
14	365	12.8	520	US-10-114-270-60	Sequence 60, Appl
15	353	12.4	524	US-10-424-599-199559	Sequence 199559,

16	352	12.3	533	12	US-10-424-599-269344	Sequence 269344,
17	350	12.3	520	12	US-10-072-012-548	Sequence 548, App
18	350	12.3	520	14	US-10-032-189-115	Sequence 115, App
19	349.5	12.3	520	12	US-10-424-599-228819	Sequence 228819,
20	349	12.2	508	14	US-10-274-694-16	Sequence 16, Appl
21	348.5	12.2	546	15	US-10-291-265-405	Sequence 405, App
22	347.5	12.2	524	15	US-10-291-265-877	Sequence 877, App
23	347.5	12.2	524	15	US-10-291-265-878	Sequence 878, App
24	347.5	12.2	524	15	US-10-291-265-879	Sequence 879, App
25	346.5	12.1	524	9	US-09-989-722-264	Sequence 264, App
26	346.5	12.1	524	9	US-09-989-723-264	Sequence 264, App
27	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
28	346.5	12.1	524	9	US-09-989-731-264	Sequence 264, App
29	346.5	12.1	524	9	US-09-989-732-264	Sequence 264, App
30	346.5	12.1	524	9	US-09-991-073-264	Sequence 264, App
31	346.5	12.1	524	9	US-09-990-442-264	Sequence 264, App
32	346.5	12.1	524	9	US-09-991-163-264	Sequence 264, App
33	346.5	12.1	524	9	US-09-993-604-264	Sequence 264, App
34	346.5	12.1	524	9	US-09-990-456-264	Sequence 264, App
35	346.5	12.1	524	9	US-09-989-721-264	Sequence 264, App
36	346.5	12.1	524	9	US-09-989-735-264	Sequence 264, App
37	346.5	12.1	524	9	US-09-989-735-264	Sequence 264, App
38	346.5	12.1	524	9	US-09-989-735-264	Sequence 264, App
39	346.5	12.1	524	9	US-09-990-444-264	Sequence 264, App
40	346.5	12.1	524	9	US-09-991-181-264	Sequence 264, App
41	346.5	12.1	524	9	US-09-989-730-264	Sequence 264, App
42	346.5	12.1	524	9	US-09-990-436-264	Sequence 264, App
43	346.5	12.1	524	9	US-09-993-687-264	Sequence 264, App
44	346.5	12.1	524	10	US-09-989-734-264	Sequence 264, App
45	346.5	12.1	524	10	US-09-989-734-264	Sequence 264, App

ALIGNMENTS

RESULT 1

US-10-066-007-1
Sequence 1, Application US/10066007
Publication No. US2003007691A1
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: OJIMA, Kazuyuki
APPLICANT: SETOGUCHI, Yutaka
TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
FILE REFERENCE: ASTAXANTHIN SYNTHETASE
CURRENT APPLICATION NUMBER: US/10/066,007
CURRENT FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: US/09/518,386
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: EP 99104668.1
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: EP 00101666.6
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 557
TYPE: PRT
ORGANISM: Phaffia rhodozyma
FEATURE:
NAME/KEY: TRANSIT
LOCATION: (1)..(26)
US-10-066-007-1

Query Match 100.0%; Score 2852; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.3e-248;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFTLVLLTGALGLAFAFWSASIAFFSLYLAPRRSSLYNLQGNHNTYFTCNFLDLSARTG 60
Db 1 MFTLVLLTGALGLAFAFWSASIAFFSLYLAPRRSSLYNLQGNHNTYFTCNFLDLSARTG 60
Qy 61 EEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFVNHVMEAYDYPKPGMAARVLRIATGDG 120

```
Db 61 EEHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRATG 120
QY 121 VVTAGEAHRHRRIMISLSQAQVKSMPVIFLEKGMELVDKMWEDAAEKDVAVGESAGE 180
Db 121 VVTAGEAHRHRRIMISLSQAQVKSMPVIFLEKGMELVDKMWEDAAEKDVAVGESAGE 180
QY 181 KKAATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDGFAPTL 240
Db 181 KKAATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDGFAPTL 240
QY 241 SFKAIMWDFVPFRFMKRRHEIPLTOGLAVSRVGVIELMEQKQAVLGSASDAQVDDKDV 300
Db 241 SFKAIMWDFVPFRFMKRRHEIPLTOGLAVSRVGVIELMEQKQAVLGSASDAQVDDKDV 300
QY 301 QGRDILSLVRANIAANLPEQKLSDEEVLQAINLFAGYETSTVLTWFMFRLSEDKA 360
Db 301 QGRDILSLVRANIAANLPEQKLSDEEVLQAINLFAGYETSTVLTWFMFRLSEDKA 360
QY 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLSRLDPPSPYANRECLKDEDFTPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLSRLDPPSPYANRECLKDEDFTPLAE 420
QY 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFYIGEDAEEFRERWLEDVTDLSNSIEA 480
Db 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFYIGEDAEEFRERWLEDVTDLSNSIEA 480
QY 481 PYGHOASFISSGPRACFGWRFAVAKMFLVTLRRVQFEPHSHPEYEHITLIISRPRIV 540
Db 481 PYGHOASFISSGPRACFGWRFAVAKMFLVTLRRVQFEPHSHPEYEHITLIISRPRIV 540
QY 541 GREKEGYQMLQVKEPVE 557
Db 541 GREKEGYQMLQVKEPVE 557
```

RESULT 2

```
US-10-066-007-3
; Sequence 3, Application US/10066007
; Publication No. US20030077691A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518,386
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US/09/518,386
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-10-066-007-3
```

```
Query Match 100.0%; Score 2852; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.3e-248;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFTVLVLTGALGLAAPSWSIASIAFFSLYLAPRESSLYNLOGNHTNYFTGNFLDILSARTG 60
Db 1 MFTVLVLTGALGLAAPSWSIASIAFFSLYLAPRESSLYNLOGNHTNYFTGNFLDILSARTG 60
QY 61 EEHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRATG 120
Db 61 EEHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRATG 120
```

```
QY 121 VVTAGEAHRHRRIMISLSQAQVKSMPVIFLEKGMELVDKMWEDAAEKDVAVGESAGE 180
Db 121 VVTAGEAHRHRRIMISLSQAQVKSMPVIFLEKGMELVDKMWEDAAEKDVAVGESAGE 180
QY 181 KKAATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDGFAPTL 240
Db 181 KKAATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDGFAPTL 240
QY 241 SFKAIMWDFVPFRFMKRRHEIPLTOGLAVSRVGVIELMEQKQAVLGSASDAQVDDKDV 300
Db 241 SFKAIMWDFVPFRFMKRRHEIPLTOGLAVSRVGVIELMEQKQAVLGSASDAQVDDKDV 300
QY 301 QGRDILSLVRANIAANLPEQKLSDEEVLQAINLFAGYETSTVLTWFMFRLSEDKA 360
Db 301 QGRDILSLVRANIAANLPEQKLSDEEVLQAINLFAGYETSTVLTWFMFRLSEDKA 360
QY 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLSRLDPPSPYANRECLKDEDFTPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLNAALPYLEAFVKSLSRLDPPSPYANRECLKDEDFTPLAE 420
QY 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFYIGEDAEEFRERWLEDVTDLSNSIEA 480
Db 421 PVIGRDSGVINEVRITKGTWMLPLFNINRSKFYIGEDAEEFRERWLEDVTDLSNSIEA 480
QY 481 PYGHOASFISSGPRACFGWRFAVAKMFLVTLRRVQFEPHSHPEYEHITLIISRPRIV 540
Db 481 PYGHOASFISSGPRACFGWRFAVAKMFLVTLRRVQFEPHSHPEYEHITLIISRPRIV 540
QY 541 GREKEGYQMLQVKEPVE 557
Db 541 GREKEGYQMLQVKEPVE 557
```

RESULT 3

```
US-10-313-963A-56
; Sequence 56, Application US/10313963A
; Publication No. US20040002078A1
; GENERAL INFORMATION:
; APPLICANT: Boutell, Jonathan
; APPLICANT: Godber, Benjamin
; APPLICANT: Hart, Darren
; APPLICANT: Blackburn, Jonathan
; TITLE OF INVENTION: Arrays
; FILE REFERENCE: KIL-001
; CURRENT APPLICATION NUMBER: US/10/313,963A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/335,806
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/410,815
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 56
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-313-963A-56
```

```
Query Match 14.0%; Score 398.5; DB 15; Length 503;
Best Local Similarity 28.7%; Pred. No. 6.2e-27;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAPSW--ASTAFPSLYLAPRRS--SLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLLAVALSVLLVLYGYTHSHGLFKLGIPOPTPLPFLGN---ILSYHKGFCMFDM 63
QY 62 BHAKYREKYGSTLRAGAPVLNSTDPKVFNVHMKAEYDYPKPGMAARVLRAT 117
Db 64 CHKYGKVGWGF---YDG--QQPVLAITDPMIKTVLVKECYSVFTNRRPFGVFKSA- 117
QY 118 GDGVVTAEGAEHRRHRRIMISLSQAQVKSMPVIFLEKGMELVDKMWEDAAEKDVAVGES 177
```

Db 118 ----ISAEDEEWKRLSLLSPFTSGKLEMPVPIIAQGDVULVNLREA----- 164
QY 178 AGEKKAATRLT-EGVDVQKDWGRATLDVNALAGFYKSDSLQNKTNELVAVFGLTDGFA 236
Db 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNDLSLNNPD----- 201
QY 237 PTLDSFKAIM-WDFV-PYFRMK-RRHEIPLTQGLAV---SRVGIEMEQKQAVLGS 290
Db 202 PFVENTKLLRFDLDFLFLSITVFFFLIPIILEVNI CVFPREVTNFRKSVKR-----M 256
QY 291 SDQAVDKVQGRDILSLAVRANIAANLPESOKLSDEEVLAQISNLLFAGYETSTVLTW 350
Db 257 KESRLEDTQKRVDFLQIMDSQKTESHKALSDELVAQSIIFIFAGYETTSVLSF 316
QY 351 MFHRLSEDAVQKLEBEICQI---DTDMPTLDLNLPLFNINRSKFIYGEDAEERPERWL 408
Db 317 IMVELATHPDVQKLEBEIDAVLPNKAPPTYDTVLOMEYLVMMVNETLRLFPFIARLERV 376
QY 409 CLKDEDFIPLAEPVIGRDSVINEVRITKGTMMVPLFNINRSKFIYGEDAEERPERWL 468
Db 377 CKKDE-----INGMFIKGVVWVIPSVALHRDP-KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWFAVAEMKAPFLVTLRRVQPEP 520
Db 421 KKKKNIDPIYITPFG-----SGPRNCIGMRPALMMKALIRVLQNFSPKP 467

RESULT 4

US-10-146-575-2
; Sequence 2, Application US/10146575
; Publication No. US20030059800A1
; GENERAL INFORMATION:
; APPLICANT: Lichter, Jay
; APPLICANT: Guido, Marco
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-12P
; CURRENT APPLICATION NUMBER: US/10/146,575
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US/09/144,367
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 503
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-146-575-2

Query Match 13.9%; Score 396.5; DB 14; Length 503;
Best Local Similarity 28.7%; Pred. No. 9.4e-27;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
QY 13 LAAFSW--ASIAFFSYLAPRRS--SLYNLQ--PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLAVSLVLLYLGTHSHGLFKKLGIPGPTPLPFLGN---ILSYHKGFCWFMDE 63
QY 62 EHAKYREKYGSTLRFAGIAGAPVNLSTDPKVFNHVM-KEAYDY---PKPGMAARVLRIAT 117
Db 64 CHKIKGVWGF---YDG--QOPVLAITDPMIKTVLVKECYSVFTNRRFPFGVGFKA- 117
QY 118 GDGVVTAEGEAHKKHRRIMIPISLAQAVKSMVPIFLEKGMELVDKXMEDAAEKMAVES 177
Db 118 ---ISAEDEEWKRLSLLSPFTSGKLEMPVPIIAQGDVULVNLREA----- 164
QY 178 AGEKKAATRLT-EGVDVQKDWGRATLDVNALAGFYKSDSLQNKTNELVAVFGLTDGFA 236
Db 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNDLSLNNPD----- 201
QY 237 PTLDSFKAIM-WDFV-PYFRMK-RRHEIPLTQGLAV---SRVGIEMEQKQAVLGS 290
Db 202 PFVENTKLLRFDLDFLFLSITVFFFLIPIILEVNI CVFPREVTNFRKSVKR-----M 256
QY 291 SDQAVDKVQGRDILSLAVRANIAANLPESOKLSDEEVLAQISNLLFAGYETSTVLTW 350

Db 257 KESRLEDTQKRVDFLQIMDSQKTESHKALSDELVAQSIIFIFAGYETTSVLSF 316
QY 351 MFHRLSEDAVQKLEBEICQI---DTDMPTLDLNLPLFNINRSKFIYGEDAEERPERWL 408
Db 317 IMVELATHPDVQKLEBEIDAVLPNKAPPTYDTVLOMEYLVMMVNETLRLFPFIARLERV 376
QY 409 CLKDEDFIPLAEPVIGRDSVINEVRITKGTMMVPLFNINRSKFIYGEDAEERPERWL 468
Db 377 CKKDE-----INGMFIKGVVWVIPSVALHRDP-KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWFAVAEMKAPFLVTLRRVQPEP 520
Db 421 KKKKNIDPIYITPFG-----SGPRNCIGMRPALMMKALIRVLQNFSPKP 467

RESULT 5
US-09-957-997-3
; Sequence 3, Application US/09957997
; Patent No. US20020150915A1
; GENERAL INFORMATION:
; APPLICANT: Berkenstam, Anders
; APPLICANT: Bertilsson, Gran
; APPLICANT: Blomquist, Patrik
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-046001
; CURRENT APPLICATION NUMBER: US/09/957,997
; EARLIER FILING DATE: 2001-09-21
; EARLIER APPLICATION NUMBER: SE 0003393-6
; EARLIER FILING DATE: 2000-09-22
; EARLIER APPLICATION NUMBER: 60/238,895
; EARLIER FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-957-997-3

Query Match 13.5%; Score 385.5; DB 9; Length 503;
Best Local Similarity 29.0%; Pred. No. 9.3e-26;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;
QY 13 LAAFSW--ASIAFFSYL-APRESSLYNLQ--PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAVETWLLAVSLVLLYLGTHGLFKKLGIPGPTPLPFLGN---ALSPKGYWTFDME 63
QY 62 EHAKYREKYGSTLRFAGIAGAPVNLSTDPKVFNHVM-KEAYD-----YKPGMAARVL 113
Db 64 CYKRYKVGW-----YDCCQOPMLAITDPMIKTVLVKECYSVFTNRRFPFGVGFKNAI 118
QY 114 RIATGDGVVTAEGEAHKKHRRIMIPISLAQAVKSMVPIFLEKGMELVDKXMEDAAEKDMA 173
Db 119 SI-----ADEDEWKRIIRSLSPFTSGKLEKMPVPIIAQGDVULVNLREA----- 164
QY 174 VGESAGEKKAATRLT-EGVDVQKDWGRATLDVNALAGFYKSDSLQNKTNELVAVFGLT 232
Db 165 -----ETGKPVTLKGVFGAYSMVDITSTSGVSDLSLNNPD-----PFVENT 207
QY 233 D---GPAPTLDSFKAIMWDFVYFRMKRRIEIPLTQGLAVS---RRVGIEMEQKQAV 286
Db 208 KKLRENP-LDPFVLSIKVP-PFL-----TPILEALNITVFRKVISFLTSKVKQIK 257
QY 287 LGSASQAVDKKDVGR--DILSLVRAANIAANLPESOKLSDEEVLAQISNLLFAGYETS 344
Db 258 EDEL-----KETQKRVDFLQIMDSQKSDSKTHKALSDELVAQSIIFIFAGYETT 310
QY 345 STVLTWFMHRLSEDAVQKLEBEICQIDTM-----PTLDELNALPYLEAFVRESRLD 399
Db 311 SSVLSFIIYELATHPDVQKQVQKE---IDTLPNKAPPTYDTVLOMEYLVMMVNETLRLF 367
QY 400 PPSPIANRECLNDEDFIPLAEPVIGRDSVINEVRITKGTMMVPLFNINRSKFIYGEDA 459

Db 368 PYAMBLERCKDVE-----INGMFIPKGVVMPISVVLHDP-KYWTPE 411
QY 460 EEPFRPERWLEDVDSINS-IEAPYCHOASFSISGRACGFEPAVAEMKAEFLVTLRQVF 518
Db 412 EKFLPERFSKKNKNDIPVITPFG-----SGPRNCIGMRFALVNNKIALVRVLQNFSP 465
QY 519 EP 520
Db 466 KP 467
RESULT 6
US-10-425-114-38180
; Sequence 38180, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38180
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-017-B5_FLI_1.ppe
US-10-425-114-38180

Query Match 13.5%; Score 384; DB 12; Length 537;
Best Local Similarity 27.4%; Pred. No. 1.4e-25;
Matches 162; Conservative 79; Mismatches 230; Indels 120; Gaps 24;
QY 4 LVLLTGALGAAPSWASIAFFS-LYIAPRR-SSLVNLQGNHTNY--FTGNFLDILSART 59
Db 23 LLIVYGVLG-ALLLWKAARLLRLMWEPRRLERALLAQLRGTSYRFLTG---DLREYR 78
QY 60 GESHAKYRE-----KYGSTLRFAGIAGAPVLKSTDPKVFNVHMK 98
Db 79 SKEENARPLRCHDIAGHVEFFIHGAVLEHGKTC-FWFGVPVRVTVDPDLARDWA 137
QY 99 EAY-DYPPKGMARVLRATGDGVVTAEGEAHKKRRIRIMPSLSAQAVKSMVPIFLEKGM 157
Db 138 NKFGHFEPKPPA--LTKLPSDGVANHEGKWKVKKRRILNPAFLHLEKLMPLAFSACCE 195
QY 158 ELVDKXMEDAAEKDMAGVSAAGEKATRLTEGVVDKQVKGVRATLDVMAALAGF-----D 211
Db 196 ELVSRW-----AOSLGDGSCEL-----DVPDELQTLTGDVISRFAFGSSVLEG 239
QY 212 YKSDSLQNKTNELVAFVGLTDFGAPTLDSFKAIWDFVPYFRMTKRREH-PLTQGLAVS 271
Db 240 RKIFOLQAEQERLMSII--DKFA-----VPGVMSLPTKNRMRQIKSEI 283
QY 272 RRVGIELMEQKQAVLGASDAQVDKQVGRDILSLVRANIANLPE-----QKLSDE 327
Db 284 DSILRGLIGKRMQMKQGESD-----KD-----DLGLLLESNARETGDSGGPGGLTME 334
QY 328 EVLAQISNLLFAGVETSSVLTWTFHRLSEDKAVQDKLREECQI--DTDMPTLDLNLAL 385
Db 335 EVWEECKLFVAGNETSVLLTWTWVLLSWHFEWQDBAREEVLGFGKKQPGYDGLSRL 394
QY 386 PYLEAFVKESLRLLDPPSPYANRECLDEDFIPLAEPVIGRDSGVINEVRITKGMVNLPL 445
Db 395 KTVTMILYEVLRLYPFAIFSRKTYE-----MVVGDVTPYAGVTLLELV 439

QY 446 FNINRSKIYGEDAEAEPRPERWLEDVDSINSIEA--PYGHOASFSISGRACGFEPAVA 503
Db 440 LFIHHPDIWGSDAHEPRPERFAEGVARASKORLAFPPFCW-----GPRICIGONFALL 493
QY 504 EMKAFIVTTRVQVQFPIISHPEYHITLIISPR--IVGREKEGYQMRLQ 552
Db 494 EAKMALSMILQRFQF---LAPYITHV-----PRRVIMLRPMHGAQIKLR 535
RESULT 7
US-10-425-114-67055
; Sequence 67055, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67055
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4757-032-B12_FLI_1.ppe
US-10-425-114-67055
Query Match 13.2%; Score 377.5; DB 12; Length 527;
Best Local Similarity 24.9%; Pred. No. 5.3e-25;
Matches 141; Conservative 106; Mismatches 212; Indels 107; Gaps 25;
QY 6 LLTGALGAAPSWASIAFFSL---YLAPRRSSLVNLQGNHTNYFTGNFLDILSARTG-- 60
Db 24 LLIGAL-----FFLWKPVVVTRFWRGQGGGPSY-RFLVGLSLPEIKRMAAGS 71
QY 61 -----BEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNVHMKEAYD-Y 103
Db 72 KITLDVGDHDFVPIVQPYRRWVSVDYKTFLY-WFGAVPTICVAEVLGVKQVLAETGLF 130
QY 104 PKPGMAARVLRATGDGVVTAEGEAHKKRRIRIMPSLSAQAVKSMVPIFLEKGMELVDK 163
Db 131 PKDYND-S-MEVLGKGLVLANGEDWKRREHVHPAFKPKDKLTMSVM---ADLVQRM 185
QY 164 MEDAAEKDMAGVSAAGEKATRLTEGVVDKQVKGVRATLDVMAALAGF--DYKSDSLQNK 221
Db 186 MQQMRSQ-----IQASNHEAE-IELSSEFSELTSDVIAHTAFGTSYK-----EG 229
QY 222 NELVAVFVGLTDFGAPTLDSFKAIWDFVPYFRMTKRREHPIPTQGLAVSRVRGIELMEQ 281
Db 230 KEVFA-----QKELQELTFTWLDIPAPACL-RKLKLPSTKS---SERV--BELDK 275
QY 282 KQAVLGASDAQVDKQVQ--GRDILSLVRANIANLPEKSLSDSEVLQISNLLFA 339
Db 276 KVRSLMAIIEGRLAARCTSGYGNLGLMLQAR-ALEQEGHOMLTTEEVDECKTFIFA 334
QY 340 GYETSGVLTWTFHRLSEDKAVQDKLREECQIDTD-MPTLDLNLALPYLEAFVKESLRL 398
Db 335 GQDTTSHLTLTMTFLSRYSWQHLREEVURECGDAVPNPDTVTKLKNVWVLESRL 394
QY 399 DPPSPYANRECLDEDFIPLAEPVIGRDSGVINEVRITKGMVNLPLFNINRSKIYGED 458
Db 395 YSPVVFIIR-----AVGSD--ILLRTRVPKGTMTLSIPTIALLRDKVWGQD 439
QY 459 ABEFPERWLEDVDSL-----NSTEAPYCHOASFSISGRACGFEPAVAEMKAEFLVTLR 514
Db 440 ADEFNDRPEHGVSNAAAKHPNAL-----LSFSGQFRACIGONFANLEARIGIAMILQ 492

Qy	515	RVQFE--PIISHPEYEHITLIISRP	538
Db	493	RFSFELSPNVVHAPKCAVILM---PR	515
RESULT 8			
US-10-425-114-65616			
; Sequence 65616, Application US/10425114			
; Publication No. US20040034888A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 65616			
; LENGTH: 547			
; TYPE: PRT			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: 700100628_FLI_1.p			
US-10-425-114-65616			
Query Match 13.2%; Score 376.5; DB 12; Length 547;			
Best Local Similarity 24.3%; Pred. No. 6.9e-25;			
Matches 144; Conservative 96; Mismatches 220; Indels 133; Gaps 25;			
Qy	7	LTGALGAAFSW--ASTAFFSLYLAPRSSLYNLOQPNHNY--PTGNFLDILSARTGEE	62
Db	37	LAVASVSLNLVAVTLEWAWTTPWLDRLALQAQGLKGRYELFTGLDRE--TARVNR	94
Qy	63	-----HAKYREKYSTLRPAGIAGAPVLNSTDPKVFNMKEAYD	102
Db	95	ARKKPLPLGCHDITPRVQPMHHSIKEYGK--LSFTWFGPTPRVMI DPDLVKEVLSNKF	153
Qy	103	Y--PKPCMAARVRIATGDGVTAEGEAHKHRRIMPSLSAQAVKSVVIFLEKGMEL	159
Db	154	HFGPRSSRIKRL--ANGLVNDGEGKAKHRRILNPAFHHEKIKGMWPFSTCCIEW	209
Qy	160	VDKMDAAEKDAAKDVGSAGEKKATRLTEGVVDYKDWVGRATLDVVALAGF--DYKSD--	215
Db	210	ITRW-----DNSMS--SEGSSB-----IDWPPEFQNLTDGVISRTAFGSNYQEGRR	253
Qy	216	--SLQKTNELVYAFVGLTDGFATLDSFKAIMWDFV--YFRMKRHEIPTQGLAYS	271
Db	254	IFELQELAEELI-----QSVQTFIPGWYFPLTKNNRM-----	288
Qy	272	RVVGLMEQKQAVLGSASDAQVKKDVOGR-----DILSLAVRANI--AANLPESQKLS	325
Db	289	RAIDVEIKLIRE--IIGKR-----EKDTKRETNKDDLGLLLESNTRQSGNASLIGT	341
Qy	326	DEEVLAQISNLLFAGYETSVTLTMFHRUSEDKAVQDKLREEI--CQIDTMDPTLDELNA	384
Db	342	TEDVIEECKLFYFAGMETTSVLLTWTLVLSMHPWQERAREEVLSHFGRITTPDYDSLGR	401
Qy	385	LPYLEAFVKESIRLDPSPPYANRECKDEDFIPLAEPVIGRDGVSINEVRIITKGTVMVLP	444
Db	402	LKTVTMILHEVRLPLATFTRTYKEME-----IGGIKYPAFVLELLP	446
Qy	445	LENINRSKIYGEDAEERPERWLEDVTDLSNLSIEAPYGHQASF--GPRACFGWRFA	501
Db	447	VIFIHDDPDKDASENPERF-----ANGISSATRHOAAFPFGGPRICIGQSF	499
Qy	502	VAEMKAPLFTVLRVQFE--PIISHPEYEHITLIISRPVIGREKEGYQMRLO	552
US-10-425-114-61218			
; Sequence 61218, Application US/10425114			
; Publication No. US20040034888A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 61218			
; LENGTH: 547			
; TYPE: PRT			
; ORGANISM: Oryza sativa nipponbare			
; FEATURE:			
; OTHER INFORMATION: Clone ID: LIB3431-054-B3_FLI_1.p			
US-10-425-114-61218			
Query Match 12.9%; Score 368.5; DB 12; Length 547;			
Best Local Similarity 24.6%; Pred. No. 3.6e-24;			
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;			
Qy	4	LVLTLTGALGAAFSWASIAFFS-LVLIAPRR--SSLYNLOQPNHNY--FTGNF-----	51
Db	36	ILAAAAAVALLLAVSTLEWAWTTPRLERARLAQOIRGNRYRLFTGDPENVRNR	95
Qy	52	-----LDILS-----ARTGEHAKYREKYSTLRPAGIAGAPVLNSTDPKVFN	94
Db	96	EARKKPLPLGCHDIIIPRLVPMFSKAVEHGK-----PSFTWFGPTPRVVMISDPESIR	147
Qy	95	HVKEA---YDYPKPCMAARVRIATGDGVTAEGEAHKHRRIMPSLSAQAVKSVVPI	151
Db	148	EVNSNFGHYGKPKPRLGKLL-----ASGVSYEGEKWAKHRRILNPAFHHEKIKRMLPV	203
Qy	152	FLEKGMELVDKMDAAEKDAAKDVGSAGEKKATRLTEGVVDYKDWVGRATLDVVALAGF--	210
Db	204	FSNCCTEMTRW-----ENSMSI--EGMSB-----VDWPPEFQNLTDGVISKTAFG	247
Qy	211	-DYKSD-----SLQKTNELVYAFVGLTDGFATLDSFKAIMWDFVYFRTMKRHEIPT	265
Db	248	SSYEGRRIIPQLQESAERII-----QAFRTI-----FIPGWYFPLTKNNRLR	291
Qy	266	QGLAVSRVGIEMEQKQAVLGSASDAQVKKDVOGRDILSLVLANI--AANLPESQKL	324
Db	292	E---IEREVSKLL-----RGIIKG--REKAIKNGEISNGDLLGLLVESNRESNGKALGM	342
Qy	325	SDEEVLAQISNLLFAGYETSVTLTMFHRUSEDKAVQDKLREEIC-QIDTMDPTLDELN	383
Db	343	TTDEIIECKLFYFAGMETTSVLLTWTLVLSMHPWQERAREEVLSHFGRITTPDYDSL	402
Qy	384	ALPYLEAFVKESIRLDPSPPYANRECKDEDFIPLAEPVIGRDGVSINEVRIITKGTVMWL	443
Db	403	RLKTVTMILHEVRLPPVVFTRTYKEMEL-----GGIKYPAEVT-----LML	447
Qy	444	PLFINRSKIYGEDAEERPERWLEDVTDLSNLSIEAPYGHQASFIS---GPRACFGWRF	500
Db	448	PILFIHDDPDKDASENPERFADGIGSNATK-----YQTSFPFGWGPICIGQNF	500
Qy	501	AVAEKAPLFTVLRVQFEPIISHPEYEH--ITLIISRPVIGREKEGYQMRLO	552
Db	501	ALLEAKMAICTILQRFSE---LSPSYIHAPFTVITLHP-----QHGAIKIKL	545

RESULT 10
US-10-425-114-59349
; Sequence 59349, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59349
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-OSLELIB3474019G01_FLI.pap
US-10-425-114-59349

Query Match 12.9%; Score 368.5; DB 12; Length 560;
Best Local Similarity 24.6%; Pred. No. 3.8e-24;
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLLTGALGLAAPSASIAFFS-LYLAPRR-SSLYNLOGPNHTNY--FTGNF----- 51
Db 49 ILAAAAAVALLLWAVSTLWAWTPRLRLRALRAQIGRNYRLFTGDPVENVRLNR 108
QY 52 -----LDILS-----ARTGEEHAKYREXGTLRFAGIAGAPVLNSTDPKVFN 94
Db 109 EARKKPLPGCHDIIIPVLPMSKAVEHGK-----PGFTWFGPTPRVWISDPESIR 160
QY 95 HVMKEA---YDYPKPGMAARVLIATGDGVVTAAGEAHKRRHIMIPSLSAQAVKSMVPI 151
Db 161 EVMSNKGHYGKPKPTRLGKLL-----ASGVSVYEGEKWAKHRRILNPAFHHEKIKRMLPV 216
QY 152 FLEKGMELVDKMDAEEKMAVGESAGEKATKATLETEGVDVVDKVGWGRATLDVVALAGF- 210
Db 217 FSNCTETWTRW-----ENSMIS-EGMSE-----VDVWPEFQNLTGDISKTAFG 260
QY 211 -DYKSD-----SLQNKTNELYAVFGLTDGFAPTLDSFKAIMWDFVPYFRMTKRRHEIPT 265
Db 261 SSYEGRRIIFQQAESAERII-----QAFRTI-----FIPGYWFLPTKNNRLR 304
QY 266 QGLAVSRVGIEMLEQKQAVLGSASDAQVKKDQVGRDILSLVRANI-AANLPESOKL 324
Db 305 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLDGLLVESNMRESNGKAEI 355
QY 325 SDEVLAQISNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREEIC-QIDTDMPTLDELN 383
Db 356 TTDEIIECKLFYFAGMETTSVLLTTLVLSMHPWQERAREEVLHFGRTTPDYDLS 415
QY 384 ALPYLEAFVKESLRLDPPSPYANRECLKDEDFTPLAEPVIGRDSVINEVRITKGTVMVL 443
Db 416 RLKIVTMLIYEVRLPVVFLTRTYKEMEL-----GGIKYPAEVT-----LML 460
QY 444 PLFNINRSKFIYGEDAEERFERWLEDVDTSLNSIEAPYGHQASFTIS---GPRACFGWR 500
Db 461 PILFIHDPDINGKADGEFNPGRFADGISNATK-----YQTSFPFGWGPICIGQNF 513
QY 501 AVAEMKAFVTLRRVQFPEIISHPEYH---ITLIISRPRIVGREKEGYQWRLO 552
Db 514 ALLEAKMAICTILQRFSE---LSPSYTHAPFTVITLHP-----OHGAQIKLK 558

RESULT 11
US-10-425-114-59350
; Sequence 59350, Application US/10425114

Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59350
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSFLCYP143F07_FLI.pap
US-10-425-114-59350

Query Match 12.9%; Score 368.5; DB 12; Length 562;
Best Local Similarity 24.6%; Pred. No. 3.8e-24;
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLLTGALGLAAPSASIAFFS-LYLAPRR-SSLYNLOGPNHTNY--FTGNF----- 51
Db 51 ILAAAAAVALLLWAVSTLWAWTPRLRLRALRAQIGRNYRLFTGDPVENVRLNR 110
QY 52 -----LDILS-----ASTGEEHAKYREKYSTLRFAGIAGAPVLNSTDPKVFN 94
Db 111 EARKKPLPGCHDIIIPVLPMSKAVEHGK-----PSFTWFGPTPRVWISDPESIR 162
QY 95 HVMKEA---YDYPKPGMAARVLIATGDGVVTAAGEAHKRRHIMIPSLSAQAVKSMVPI 151
Db 163 EVMSNKGHYGKPKPTRLGKLL-----ASGVSVYEGEKWAKHRRILNPAFHHEKIKRMLPV 218
QY 152 FLEKGMELVDKMDAEEKMAVGESAGEKATKATLETEGVDVVDKVGWGRATLDVVALAGF- 210
Db 219 FSNCTETWTRW-----ENSMIS-EGMSE-----VDVWPEFQNLTGDISKTAFG 262
QY 211 -DYKSD-----SLQNKTNELYAVFGLTDGFAPTLDSFKAIMWDFVPYFRMTKRRHEIPT 265
Db 263 SSYEGRRIIFQQAESAERII-----QAFRTI-----FIPGYWFLPTKNNRLR 306
QY 266 QGLAVSRVGIEMLEQKQAVLGSASDAQVKKDQVGRDILSLVRANI-AANLPESOKL 324
Db 307 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLDGLLVESNMRESNGKAEI 357
QY 325 SDEVLAQISNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREEIC-QIDTDMPTLDELN 383
Db 358 TTDEIIECKLFYFAGMETTSVLLTTLVLSMHPWQERAREEVLHFGRTTPDYDLS 417
QY 384 ALPYLEAFVKESLRLDPPSPYANRECLKDEDFTPLAEPVIGRDSVINEVRITKGTVMVL 443
Db 418 RLKIVTMLIYEVRLPVVFLTRTYKEMEL-----GGIKYPAEVT-----LML 462
QY 444 PLFNINRSKFIYGEDAEERFERWLEDVDTSLNSIEAPYGHQASFTIS---GPRACFGWR 500
Db 463 PILFIHDPDINGKADGEFNPGRFADGISNATK-----YQTSFPFGWGPICIGQNF 515
QY 501 AVAEMKAFVTLRRVQFPEIISHPEYH---ITLIISRPRIVGREKEGYQWRLO 552
Db 516 ALLEAKMAICTILQRFSE---LSPSYTHAPFTVITLHP-----OHGAQIKLK 560

RESULT 12
US-09-992-901-2
; Sequence 2, Application US/09992901
; Patent No. US20020073446A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M.

```

; APPLICANT: Chory, Joanne
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
; FILE REFERENCE: SALKINS.024DV1
; CURRENT APPLICATION NUMBER: US/09/992,901
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/527,073
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170,931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172,832
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-992-901-2

```

```

Query Match 12.9%; Score 367; DB 9; Length 520;
Best Local Similarity 22.2%; Pred. No. 4.6e-24;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

QY 4 LVLLTGALGLAFAFWASTAFESLYLAPRR-----SSLYNLOQPNHNTVFTGNFLDILS--- 56
DB 13 VLVLSVILSVIVGMSL-----LWVRPKTEEHFSGKIGPPY-HFFGNVVELVGMWL 67
QY 57 -----ARTGEEHAKYREXYGSLTRFAGIAGAPVLNST--DPKVFNVHMKEA 100
DB 68 KASSHEMPFPHNLPVLVSFVHWRKIVGATFL---VWFGTFLTRVADPDILREIFSKS 124
QY 101 VDYKPGHAAVRLRATGCVVTAEGEAHKEHREIMIPSLSAQAVKSMVPIFELEKGMELV 160
DB 125 EFYEK-NEAHLPLVQLGEGDGLSLKGEKWAHRRKILISFTHEMNLKLVFVLKSVTDWV 183
QY 161 DQWMDAAEKDMVAGESAGEKATRLTEGVVDVXDWVGRATLDYMALAGFDYKSDSLQNK 220
DB 184 DKWSKLSN-----GEVE-----VDVVEWQILTEDVISRTAFGSSE----- 222
QY 221 TNELVAVFGLTDGAP-----TLDSFKAIMDFVPYPTMKRHEIPLTOGLAV 270
DB 223 -----DGRAVRLQAQMLCAEAFQVY---PIPGYRFP-----TRGNLK 260
QY 271 SRRVG-----IELMEOKQAVLGSASDOAVDKKQVQGRDILSLVRANIAANLPESQK 323
DB 261 SKLDEKTEKSKLLKIERRQNAIDGEGEC---KEPAKDLGLMIQA-----KN 308
QY 324 LSDEVLQAISNLLFAGYETSTVLTMFHLSEKAVQDKLREICQI---DTDMPTLDE 381
DB 309 VTVQDIVEECKSFPAKQTTSNLLTWTLLSMHPWQAKARDEVLRVCGSRDVPKCH 368
QY 382 LNALPYLEAFVKESRLDPPSPYANRECLKDEDPIPLAEPVIGRDSGVINEVITKTMV 441
DB 369 VVKLTSLMILNESRLPPIVATIRAKSDVK-----LGYKIPGCTEL 413
QY 442 MLPLFNINRSKFYGEDAEFRPERWLDVTDLSNLSIAPYGHQASFSIS---GPRACFGW 498
DB 414 LIPILAVHDDQAIWGNVNEFNPAFADGVPRAAK-----HPVGPFPGLGVRTCIGQ 466
QY 499 RFAVEMKAFVTLRRVQFEPFIISHPEYEHITLII 534
DB 467 NLAILQAKLTIAVNIQRETFH---LAPTYQHAPTVL 499

```

```

RESULT 13
US-10-425-114-38827
; Sequence 38827, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

```

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38827
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700045236_FLI.pap
US-10-425-114-38827

```

```

Query Match 12.8%; Score 365.5; DB 12; Length 662;
Best Local Similarity 22.9%; Pred. No. 9.1e-24;
Matches 123; Conservative 96; Mismatches 193; Indels 125; Gaps 16;

QY 70 YGSLRPAAGTAGAPVLNSTDPKVFNVHMKYDYKPGMAARVLAIRIATGCVVTAEGEAH 129
DB 188 YGGIFRL-NFGPKSFILVSDPAIAGHILRENSKAYSKGILAEILFEVMTGTGLIPADGEIW 246
QY 130 KRURRIMPISLSAQAVKSMVPIFELEKGMELVDKMMEDAAEKDMVAGESAGEKATRLTE 189
DB 247 RVRRRAIVPALHKKYVTAMIGLFGEASQRLCEKL-----DKAA--VDGE 288
QY 190 GVDVKOWGRATLDYMALAGFDYKSDSLQNKYNELVAVFGLTDGFAPLDSFKAIMWDF 249
DB 289 DMEMESLFSRLTLDVTKAVFNVDYDLSYDNGIVEAVVTVIRE--AEMRSTSPITW-- 344
QY 250 VPYPTMKRHEIPLTOGLAVSRVGVIELMEOKQAVLGSASDQ-----AVDKKDVQG 302
DB 345 -----EPIWKDISPROKTYNEALK-----LINSTIDELIAICKRLVEQEDLQF 388
QY 303 RD-----ILSLVRANIAANLPESOKLSDEEVLQAIISNLLFAGYETSTVLTWMPH 353
DB 389 HEYMEQDPSILHFL-----LASGDDVSKQRLDMLTMLIAGHETSAVUTWTFY 440
QY 354 RLSEDKAVQDKLREICQIDTD-MPTLDLNLALPYLEAFVKESRLDPPSPYANRECLKD 412
DB 441 LLSKPKVMKQLEADSVLGDGLPIEDVKLKYTRVINESLRLYPQPPVLRSLSD 500
QY 413 EDFIPLAEPVIGRDSGVINEVITKGMVMLPLFNINRSKFYIGEDAEFRPERWLDV 472
DB 501 D-----ILGGYPIGRGEDIFISVWNLHHC9-KHWDDBAEVFNPERWPLDGP 544
QY 473 DSLNSIAPYGHQASFSISGPRACFGWRFAVEMKAFVTLRRVQF----- 519
DB 545 NP-NEINQNSY-LPFGGGRKCVGDMFATFEVTVATMLVKRFDQFQMAPGAPPVDMWTG 602
QY 520 -----PIISHPEYEHIT-----LIISRPRIVGREKEG 546
DB 603 ATIHTEGLKMTVTRRTPPVIPNLEMKIITDSQESTLSAPSVMWVAASVASGEDQG 659

```

```

RESULT 14
US-10-114-270-60
; Sequence 60, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spyttek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.

```

```

APPLICANT: Li, Li
APPLICANT: Vernet, Corine
APPLICANT: Zerhusen, Bryan D.
APPLICANT: Gorman, Linda
APPLICANT: Shenoy, Suresh G.
APPLICANT: Pena, Carol E.A.
APPLICANT: Smithson, Glenna
APPLICANT: Burgess, Catherine E.
APPLICANT: Gerlach, Valerie
APPLICANT: Padigaru, Muraidhara
APPLICANT: Shimkets, Richard A.
APPLICANT: Gangolli, Esha A.
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Casman, Stacie J.
APPLICANT: Ji, Weizhen
APPLICANT: Anderson, David W.
APPLICANT: Lieste, Mario W.
APPLICANT: Rastelli, Luca
APPLICANT: Edinger, Shomit R.
APPLICANT: Stone, David J.
APPLICANT: MacDougall, John R.
APPLICANT: Rothenberg, Mark E.
TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-322C
CURRENT APPLICATION NUMBER: US/10/114,270
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: 60/281,086
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,136
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/281,863
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/281,906
PRIOR FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/282,020
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,930
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/282,934
PRIOR FILING DATE: 2001-04-10
PRIOR APPLICATION NUMBER: 60/283,512
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/283,710
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/284,234
PRIOR FILING DATE: 2001-04-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 470
SEQ ID NO 60
LENGTH: 520
TYPE: PRT
ORGANISM: Homo sapiens
US-10-114-270-60

Query Match 12.8%; Score 365; DB 12; Length 520;
Best Local Similarity 24.6%; Pred. No. 6.9e-24;
Matches 143; Conservative 104; Mismatches 221; Indels 114; Gaps 23;

QY 2 FILVLLTGNLGLAA--PSWASIAFSLYLAPRSSLLNLO---GPNHTYVFTGNFLDILS 56
Db 19 WILLLLVGSWLLAILAWT-----YTFY----DNCRLQCFQPPQKQWFWGHLGLVTP 69
QY 57 ARTGEEH-----AKYREKY-----GSTLRFAGIAGAPVLNSTDPKVFNVHMKDAYDYPKPG 107
Db 70 TEEGMKTLTQLVATVPQGFVKWLGPIIFVLCHPDITRSI-----TWASAAIVPKDN 122
QY 108 MAAPVLRIATGCVVTAGEAHKRRIMTSPLSAQAVKSMVPIFELEKMGELVDKMWEDA 167
Db 123 LFYSFLKPLWGLGLSLGSGDKSWRRHRLMTAPAFHNLKSYIIFKKSANIMLDKQWHLA 182
QY 168 AEKDMAVGSAGESAKKXATRIETEGVDVKDWVGRATLDWMALAGFDYKDSLQNTNELYVA 227
Db 183 SE-----GSSR-----LDMPEHISLMTLDSLOKCVSFSES-NCQSKPSE-YIA 223

```

QY	268	LAVSRVGIEMEQKQAVLGSASDAQVDKXQVQGRDILSLVRA-----NIA	315
Db	265	---DRYIKASLTDMIKK-----REKAPKTGEATRDLLGILLESNHKEIQEHRNNVNG	315
QY	316	ANLPESQKLSDEEVLQAISNLLFAGYETSTVLTWTFHRLSEDKAVQDKLREEICQI-DT	374
Db	316	NNL-----NDVIEECKLYFAGQETTSVLLVWTVLLSRYPDWQSRAREEVLOVFGK	367
QY	375	DMPTLDELNALPYLE-AFVKESLRDPPSPYANRECLKDEDFIPLABPVGIRGDSVINEV	433
Db	368	QAPNFDGLSHLKIVTMIFLKVLRLYPYPAVGLNEN-----VDRDMKLG-N-L	412
QY	434	RITKGTWVMLPLFNINRSKFIYGEDAEFFRPERWLEDVTDLSNLSIA--PYGHQASFI	491
Db	413	SLPAGVQVSLPTTVPVPHDELWGDVNEFFPERFSEGVKATNGRVSPFFFGW-----G	466
QY	492	PRACFGWRPVAEMKAFLEVTIRRVQPEPIISHPEYEH--ITLII	538
Db	467	PRICIGQNFSLLEAKQALSTILOHFSFE---LSPAYAHAFVTVFTLQPQ	512

Search completed: April 2, 2004, 14:10:36
 Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17.5 Seconds
(without alignments)
1643.181 Million cell updates/sec

Title: US-10-066-007-1
Perfect score: 2852
Sequence: 1 MFILVLLTGALGLAFAFWSAS.....RIVGREKEGVQKRLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- 1: Issued Patents AA.*
- 2: /cgn2_6/ptodata/2/iaa/5A COMB pep.*
- 3: /cgn2_6/ptodata/2/iaa/5B COMB pep.*
- 4: /cgn2_6/ptodata/2/iaa/6A COMB pep.*
- 5: /cgn2_6/ptodata/2/iaa/6B COMB pep.*
- 6: /cgn2_6/ptodata/2/iaa/PTCUS COMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	4	US-09-518-386B-1
2	2852	100.0	557	4	US-09-518-386B-3
3	396.5	13.9	503	4	US-09-144-357-2
4	370	13.0	504	1	US-08-457-274A-25
5	370	13.0	504	5	PCT-US95-05758-25
6	367	12.9	520	4	US-09-527-073-2
7	345	12.1	524	4	US-09-976-594-533
8	338.5	11.9	503	4	US-09-583-447A-2
9	331	11.6	540	4	US-09-302-620B-99
10	330.5	11.6	507	1	US-08-457-274A-22
11	330.5	11.6	507	5	PCT-US95-05758-22
12	328	11.5	504	4	US-09-583-447A-4
13	328	11.5	540	4	US-09-302-620B-98
14	326.5	11.4	541	3	US-09-158-767-19
15	326.5	11.4	541	3	US-09-158-767-20
16	321	11.3	576	3	US-08-948-564-16
17	318.5	11.2	526	1	US-08-298-426-4
18	294.5	10.3	489	4	US-09-852-067-4
19	294	10.3	522	4	US-09-302-620B-97
20	289	10.1	522	4	US-09-302-620B-96
21	276	9.7	420	4	US-09-583-447A-6
22	276	9.7	467	4	US-09-126-420A-17
23	269.5	9.4	517	4	US-09-302-620B-100
24	269.5	9.4	517	4	US-09-911-781-32
25	269	9.4	507	1	PCT-US95-05758-23
26	269	9.4	507	5	PCT-US95-05758-23
27	267.5	9.4	517	4	US-09-302-620B-101

28	263	9.2	523	4	US-09-302-620B-95	Sequence 95, Appl
29	259.5	9.1	510	3	US-08-948-564-4	Sequence 4, Appl
30	259	9.1	510	4	US-09-852-067-2	Sequence 2, Appl
31	251.5	8.8	508	4	US-09-126-420A-25	Sequence 25, Appl
32	246	8.6	512	4	US-09-302-620B-103	Sequence 103, Appl
33	243	8.5	512	4	US-09-302-620B-102	Sequence 102, Appl
34	241.5	8.5	476	1	US-08-333-075A-30	Sequence 30, Appl
35	236.5	8.3	498	1	US-08-457-274A-24	Sequence 24, Appl
36	236.5	8.3	498	5	PCT-US95-05758-24	Sequence 24, Appl
37	235.5	8.3	499	4	US-09-302-620B-104	Sequence 104, Appl
38	235.5	8.3	504	4	US-09-126-420A-18	Sequence 18, Appl
39	233.5	8.2	509	4	US-09-499-302A-8	Sequence 8, Appl
40	228	8.0	426	2	US-08-560-398-4	Sequence 4, Appl
41	227	8.0	504	4	US-09-499-302A-6	Sequence 6, Appl
42	226.5	7.9	500	3	US-09-292-768-68	Sequence 68, Appl
43	226.5	7.9	500	3	US-09-292-768-70	Sequence 70, Appl
44	225.5	7.9	500	3	US-09-292-768-4	Sequence 4, Appl
45	223.5	7.8	524	4	US-09-126-420A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-518-386B-1
; Sequence 1, Application US/09518386B
; Patent No. 6365386
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518.386B
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: TRANSIT
; LOCATION: (1)..(26)
US-09-518-386B-1

Query Match 100.0%; Score 2852; DB 4; Length 557;
Best Local Similarity 100.0%; Pred. No. 4.4e-278;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MFILVLLTGALGLAFAFWSASIAFESLYLAPRRSLYNLQGNHTNYFTGNFLDLSARTG	60
DB	1	MFILVLLTGALGLAFAFWSASIAFESLYLAPRRSLYNLQGNHTNYFTGNFLDLSARTG	60
QY	61	BEHAKYRKYGSTLRPAGIAGAPVNLSTDPKFNHVMKEAYDPKPGMAARVLRATG	120
DB	61	BEHAKYRKYGSTLRPAGIAGAPVNLSTDPKFNHVMKEAYDPKPGMAARVLRATG	120
QY	121	VVTAGEAHKRRHRRIMIPSLSAQAVKSMVPIFEKGMELVDKMDAAEKDMVAGESAGE	180
DB	121	VVTAGEAHKRRHRRIMIPSLSAQAVKSMVPIFEKGMELVDKMDAAEKDMVAGESAGE	180
QY	181	KKATLETEGVDDKDWGRATLDYMALAGFDYKSDSLQNTKLNELYVAFVGLTDGFAPTLD	240
DB	181	KKATLETEGVDDKDWGRATLDYMALAGFDYKSDSLQNTKLNELYVAFVGLTDGFAPTLD	240
QY	241	SFKAIWMDFPVYFRTMKRHEIPLTOGLAVSRVVGIELMEQKQAVLGSASDAQVKKDV	300
DB	241	SFKAIWMDFPVYFRTMKRHEIPLTOGLAVSRVVGIELMEQKQAVLGSASDAQVKKDV	300

QY 301 QGRDILSVLRANAAANLPSQKLSDEEVLAAQISNLLFAGYETSSTVLTWTFHRLSEDKA 360
 Db 301 QGRDILSVLRANAAANLPSQKLSDEEVLAAQISNLLFAGYETSSTVLTWTFHRLSEDKA 360
 QY 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420
 Db 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420
 QY 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480
 Db 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480
 QY 481 PYGHCASPIGPRACFGWRFAVAKMKAFLVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
 Db 481 PYGHCASPIGPRACFGWRFAVAKMKAFLVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
 QY 541 GREKEGYQMRLOVKPVE 557
 Db 541 GREKEGYQMRLOVKPVE 557

RESULT 2

US-09-518-386B-3
 ; Sequence 3, Application US/09518386B
 ; Patent No. 6365386
 ; GENERAL INFORMATION:
 ; APPLICANT: HOSHINO, Tatsuo
 ; APPLICANT: OJIMA, Kazuyuki
 ; APPLICANT: SETOGUCHI, Yutaka
 ; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
 ; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
 ; CURRENT APPLICATION NUMBER: US/09/518,386B
 ; CURRENT FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: EP 99104668.1
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: EP 00101666.6
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 557
 ; TYPE: PRF
 ; ORGANISM: Phaffia rhodozyma
 US-09-518-386B-3

Query Match 100.0%; Score 2852; DB 4; Length 557;
 Best Local Similarity 100.0%; Pred. No. 4.4e-278;
 Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPILVLTGALGLAASFASWASTAFPSLYLAPRRSLYNLQGNPHNTYFTGNFLDILSARTG 60
 Db 1 MPILVLTGALGLAASFASWASTAFPSLYLAPRRSLYNLQGNPHNTYFTGNFLDILSARTG 60
 QY 61 BEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFHVMKEAYDYPKPGMAARVLRATG 120
 Db 61 BEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFHVMKEAYDYPKPGMAARVLRATG 120
 QY 121 VVTAEGEAHKHRRIMIPISLQAQVKSMPFIKMGKMLVDKMDAEDAKONAVGESAGE 180
 Db 121 VVTAEGEAHKHRRIMIPISLQAQVKSMPFIKMGKMLVDKMDAEDAKONAVGESAGE 180
 QY 181 KKAATELETGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTGDGAPTLD 240
 Db 181 KKAATELETGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTGDGAPTLD 240
 QY 241 SFKALMDFVYFRTMKRHEIPIUTQGLAVSRVGVIELMEQKQAVLGASDAQVDDKDV 300
 Db 241 SFKALMDFVYFRTMKRHEIPIUTQGLAVSRVGVIELMEQKQAVLGASDAQVDDKDV 300
 QY 301 QGRDILSVLRANAAANLPSQKLSDEEVLAAQISNLLFAGYETSSTVLTWTFHRLSEDKA 360
 Db 301 QGRDILSVLRANAAANLPSQKLSDEEVLAAQISNLLFAGYETSSTVLTWTFHRLSEDKA 360

QY 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420
 Db 361 VQDKLREIEICQIDTMDPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDEDFIPLAE 420
 QY 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480
 Db 421 PVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWLEDVTDLSNSTEA 480
 QY 481 PYGHCASPIGPRACFGWRFAVAKMKAFLVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
 Db 481 PYGHCASPIGPRACFGWRFAVAKMKAFLVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
 QY 541 GREKEGYQMRLOVKPVE 557
 Db 541 GREKEGYQMRLOVKPVE 557

RESULT 3

US-09-144-367-2
 ; Sequence 2, Application US/09144367
 ; Patent No. 6432639
 ; GENERAL INFORMATION:
 ; APPLICANT: Lichter, Jay
 ; APPLICANT: Guido, Marco
 ; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
 ; FILE REFERENCE: SEQ-12P
 ; CURRENT APPLICATION NUMBER: US/09/144,367
 ; CURRENT FILING DATE: 1998-08-31
 ; PRIOR APPLICATION NUMBER: 60/058,612
 ; PRIOR FILING DATE: 1997-09-10
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 503
 ; TYPE: PRF
 ; ORGANISM: H. sapiens
 US-09-144-367-2

Query Match 13.9%; Score 396.5; DB 4; Length 503;
 Best Local Similarity 28.7%; Pred. No. 5.8e-31;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQG--PNHNY--FTGNFLDILSARTG-----E 61
 Db 7 LAMETWLLAVSLVLLYLYGTHSHGLFKLGIPTPLPFLGN--ILSYHKGFCMFDM 63
 QY 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFHVM--KEAYD--PKPGMAARVLRAT 117
 Db 64 CHKKYKVGWGF--YDG--QOPVLAITDDPMIXTLVKCYSVFTNRRPFGVGFWSA- 117
 QY 118 GDGVVTAEGEAHKHRRIMIPISLQAQVKSMPFIKMGKMLVDKMDAEDAKONAVGES 177
 Db 118 ---ISIADEDEWKLRLSLLSPTFTSGKLKEMVPIIAQYGDVLRNLREBA----- 164
 QY 178 AGEKKAATELET--EGVDYKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTGDGA 236
 Db 165 -----ETGKVTILKDVFGAYSMDVITSFGVNIIDSLNPNQD----- 201
 QY 237 PTLDSFKAIM--WDFV--PYFTMK--RRHEIPTQGLAV---SRRVGIELMEQKQAVLOSA 290
 Db 202 PFVENTKKLRFDFDLDFPFLSITVFPPILILEVLNICVFPREVTNFKRSVKR-----M 256
 QY 291 SDQAVDKDQVQGRDILSVLRANAAANLPSQKLSDEEVLAAQISNLLFAGYETSSTVLTW 350
 Db 257 KESRLEDQKHEVDFLOLMIDMSQNSKETESHKALSDELVAQSIIFIPAGYETTSVLSF 316
 QY 351 MPRHUSEDKAVQDKLREIEICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANRE 408
 Db 317 IMYELATHPDVQKQLOEIDEAVLFNKAPPTDYTVLQMSYLDVMVNETLRLFPFIAMRLERV 376
 QY 409 CLKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPRERWL 468

Db 377 CKDVE-----INGNFIKGVVWVPSYALHRDP-KYWTPEKFLPERFS 420
 QY 469 EDVTSLSN-IEAPYGHQASISGPRACFGWRPFAVEMKAFVTLRRVQFEP 520
 Db 421 KKNKNIDYIYTPFG-----SGPRNCIGNRPAALNNKMLALRLVQLNFSFXP 467

RESULT 4

US-08-457-274A-25
 ; Sequence 25, Application US/08457274A
 ; Patent No. 5734086
 ; GENERAL INFORMATION:
 ; APPLICANT: Scott, Jeffrey G.
 ; APPLICANT: Tomita, Takashi
 ; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/457,274A
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman, Michael L.
 ; REGISTRATION NUMBER: 30,727
 ; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 716-263-1304
 ; TELEFAX: 716-263-1600
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 504 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Rat
 ; STRAIN:
 ; DEVELOPMENTAL STAGE: Adult
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT:
 ; US-08-457-274A-25

Query Match 13.0%; Score 370; DB 1; Length 504;
 Best Local Similarity 26.2%; Pred. No. 2.7e-28;
 Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;
 QY 7 LTCALGLAFAFWSASIAFFSLY-LAPRRSSLYNLOQ-----PNHTNYFTGNF-LDI 54
 Db 3 LLSALTLETWLLAVLVLLYGFGRTHGLFKQIGPGKPLPFGTVLNYWGLWKFDV 62
 QY 55 LSARTGEHAKREKXGSLRAGIAGAPVLNSTDPKVFNVHM-KEAY-----DYPKP 106
 Db 63 -----ECHKYKYGKING-----LFDG--QMPLFAITDTEMIKKVLVKCFSVFTNRDFGPV 111
 QY 107 GMAARVLRIATGDGVVTAEGEAAKRRRIMIPSLSAQAVKSWPIFLEKGMELVDXOMED 166
 Db 112 GI-----MGKAVSVADEEKRYRALLSPFTTSGRLEKMPFIIEQYGDILVYLKQE 163
 QY 167 AAEKDNVAVGESAGEKATLET-EGVDVQDWVGRATLDVMAAGFDYKSDSLQN----- 219

Db 164 A-----ETGKPVTKKVFAGYSMDVITSTSGVNVDSLNNPKDPFV 204
 QY 220 -KTNEL-----YVAFVGLTDGFAPTLDSFKAIW--DFVPYFRMTKRRHEIPL 264
 Db 205 ERTKLLRFDPDFLFLSVLPFJT-----PIYMLNICHFPKDSIEFFK----- 250
 QY 265 TOGLAVSRVGIELMEQKQAVLGSASDAQVKKDVGGRDILSLVLRANLANLPSQ-K 323
 Db 251 -----KFYVRMKETRLDSVQKHRV-----DFLQLMNAHNDSDKSHSTA 290
 QY 324 LSDEEVLAQISNLLFAGYETSTVLTWMEHRLSEDKAVQDKLREEICQI--DTDMPETLDE 381
 Db 291 LSDMEITAGSIIFIFAGYEPTSSLTSLFVLHSLATHPTOKKLOEIDRALPNKAPTYDT 350
 QY 382 LNALPYLEAFVKESLRDPPSPYANRECLDKDEDFIPLAEPVIGRDGVSINEVITKGTWV 441
 Db 351 VMEMEYLDVNLNETLRLYPIGNLRLRVCKKDVE-----INGVFMFKGSVV 395
 QY 442 MLPLFNIRSKFIYGEDAEERPERWLEDVDTSLNS-IEAPYGHQASISGPRACFGWRP 500
 Db 396 MIPSYALHEDPQHWP-EERPERPERFSKENKGSIDPVYVLPFG-----NGPRNCIGMRP 448
 QY 501 AVAEMKAFVTLRRVQFEP 520
 Db 449 ALNNMKLALTQVLFQFOP 468

RESULT 5

PCT-US95-05758-25
 ; Sequence 25, Application PC/TUS9505758
 ; GENERAL INFORMATION:
 ; APPLICANT: Cornell Research Foundation, Inc.
 ; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its
 ; TITLE OF INVENTION: Uses
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
 ; STREET: P.O. Box 1051, Clinton Square
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/05758
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman, Michael L.
 ; REGISTRATION NUMBER: 30,727
 ; REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 716-263-1304
 ; TELEFAX: 716-263-1600
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 504 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Rat
 ; STRAIN: Unknown
 ; DEVELOPMENTAL STAGE: Adult
 ; POSITION IN GENOME:

CHROMOSOME/SEGMENT: Unknown

Query Match 13.0%; Score 370; DB 5; Length 504;
 Best Local Similarity 26.2%; Pred. No. 2.7e-28;
 Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY 7 LTGALGALAAFWASIAFSLY-LAPRRSSLYNQ-----PNTNPTGNF-LDI 54
 DB 3 LLSALTLETTWLLAVLLVLLVGLGTRHGLFKQIGPGKPLPFFGTVLNYGLWLMKFDV 62
 QY 55 LSARTGEHAKYRKYGSTLRFCAGAPVLNSTDPKVENHYM-KEAY-----DYKPK 106
 DB 63 -----ECHKKYKING--LFDG--QMLFALDTDEMINKVILKCEFSVFNRRDFGV 111
 QY 107 GMAARVLRIATGCVTAEGEAHKKRRIMPSLSAQAVKSMVPIFEKGMELVDMED 166
 DB 112 GI-----MKAIVSAKDEEMKRYRALLSPTTSGRIKEMFPIEQYGDILVKYLKQE 163
 QY 167 AAEKDMAYGESAGEKATRLTETEGVOVDVWGRATLDVVALAGFDYKSDSLON----- 219
 DB 164 A-----ETGKPYTMKKVFCAYSMDVITSTSGVNVDSLNPKOPFV 204
 QY 220 -KTNEL-----YVAFVGLTDGAPLTDLSFKALMW--DFVPFRVWRHREIPL 264
 DB 205 EKTGKLRFDFDFLFLSVLFPFLT-----PIYEMLNICMFKDSIEFFK----- 250
 QY 265 TOGLAVSRVVGIELMEQKQAVLGASDAQVDKQVQGRDILSLVVRANIANLPESQ-K 323
 DB 251 -----KPYRMKTRLDVQKGRV-----DFQLQMMNHNDSKOKESHTA 290
 QY 324 LSDREVLQAINLIPAGYETSSVTLTMFHRLESDKAVQDKLREICQI--DTDMPTLDE 381
 DB 291 LSDMEITAQSIIFIPAGYETSSVTLTMFHRLESDKAVQDKLREICQI--DTDMPTLDE 350
 QY 382 LNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAEFVIGRDSGVINEVRIKGTMV 441
 DB 351 VMEMEYLDMLNLTLPYIGNRLERVCKKQVE-----INGVPMKGSVV 395
 QY 442 MLPFNINRSFYIGEDAEEFRPRERWLEDVTDLSNLSIAPYGHQASFTIS---GPRACFGW 500
 DB 396 MPSTALHRDQHWPE--FEFRERFSKKNKGSIDPVYVLPFG-----NGPRNCIGNRF 448
 QY 501 AVAEMKAFVTLRRVQFEP 520
 DB 449 ALMNWKLATKVLQNFSPQ 468

RESULT 6

US-09-527-073-2
 ; Sequence 2, Application US/09527073
 ; Patent No. 6534313
 ; GENERAL INFORMATION:
 ; APPLICANT: Michael M. Neff
 ; APPLICANT: Joanne Chory
 ; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
 ; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
 ; FILE REFERENCE: SALKINS.024A
 ; CURRENT APPLICATION NUMBER: US/09/527,073
 ; PRIOR FILING DATE: 2000-03-16
 ; PRIOR APPLICATION NUMBER: US 60/124570
 ; PRIOR FILING DATE: 1999-03-16
 ; PRIOR APPLICATION NUMBER: US 60/170,931
 ; PRIOR FILING DATE: 1999-12-14
 ; PRIOR APPLICATION NUMBER: US 60/172,832
 ; PRIOR FILING DATE: 1999-12-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 520
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana

US-09-527-073-2

Query Match 12.9%; Score 367; DB 4; Length 520;
 Best Local Similarity 22.2%; Pred. No. 5.7e-28;
 Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

QY 4 LVLLTGALGALAAFWASIAFSLYLAARR-----SSLYNQGNHNTNYFTGNFLDILS--- 56
 DB 13 VLVLSVLVSVKGNLSL-----LWNRPRKTEEHFSKQIRGPPY-HFFIGNVKELVGMML 67
 QY 57 -----ARTGEHAKYRKYGSTLRFCAGAPVLNST--DPKVENHYMKEA 100
 DB 68 KASGHPMPFSHNILPVLVSFYHWRKIYGATFL--VWFGPTFRLTVADPDILREIFSKS 124
 QY 101 VDYKPGMAARVLRIATGCVTAEGEAHKKRRIMPSLSAQAVKSMVPIFEKGMELV 160
 DB 125 EYFEK-NEAHPVQLQEGDGLLSLKEKWAHKKIISPTFHENIKLLVPVVLKSVTDMV 183
 QY 161 DOMMEDAAEKDMAYGESAGEKATRLTETEGVOVDVWGRATLDVVALAGFDYKSDSLONK 220
 DB 184 DKWSKLSN-----GEVE-----VDYEMFQILTEDVISRTAFGSSYE----- 222
 QY 221 TNELYVAFVGLTDGFAP-----TLDSEKAIMWDFVYFRTMKRREHEIPLTQGLAV 270
 DB 223 -----DGRAVRLQAQOMLICEAFQKV--FIPGYRPPP-----TGNLK 260
 QY 271 SRRVG-----IELMEQKQAVLGASDAQVDKQVQGRDILSLVVRANIANLPESQK 323
 DB 261 SRKLDKEIRKSLKLIERRQNAIDGEGEC---KEPAAKDLGLMIOA-----KN 308
 QY 324 LSDREVLQAINLIPAGYETSSVTLTMFHRLESDKAVQDKLREICQI--DTDMPTLDE 381
 DB 309 VTQVDIVVECKSPFFACKQTSNLLTWILLSMHPWQAKARDEVLCVGRSDVPTKDH 368
 QY 382 LNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAEFVIGRDSGVINEVRIKGTMV 441
 DB 369 VVKLKTLSMILNLSRLYPPIVATIRAKSDVK-----LGGYKIPCGTEL 413
 QY 442 MLPFNINRSFYIGEDAEEFRPRERWLEDVTDLSNLSIAPYGHQASFTIS---GPRACFGW 498
 DB 414 LIPIAVHDDQAINGVNDVNEFNARFADGVPRAAK-----HPVGFIPFGLGVRTCIQ 466
 QY 499 RFAVEMKAFVTLRRVQFEPISHPYEHTLII 534
 DB 467 NLALQAKLTAVMIQRTFTH--LAPTYQHAPTVL 499

RESULT 7

US-09-976-594-533
 ; Sequence 533, Application US/09976594
 ; Patent No. 6673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; APPLICANT: Buchbinder, Jenny
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976,594
 ; PRIOR FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,409
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 533
 ; LENGTH: 524
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6673549 6825202CD1
 ; US-09-976-594-533

Query Match 12.1%; Score 345; DB 4; Length 524;
 Best Local Similarity 23.4%; Pred. No. 9.4e-26;
 Matches 140; Conservative 104; Mismatches 207; Indels 148; Gaps 25;

RESULT 9

RESULTS
US-09-302-620B-99
: Sequence 99, Application US/09302620B

sequence 99, application 00/035020200
; Patent No. 6331420
; GENERAL INFORMATION:

; GENERAL INFORMATION:
 ; APPLICANT: Wilson, C. Ron
 ; APPLICANT: Craft, David L.

APPLICANT: Eirich, Dudley
APPLICANT: Eshoo, Mark

APPLICANT: Madduri, Krishna M.
APPLICANT: Cornett, Cathy A.

APPLICANT: Brenner, Alfred A.
APPLICANT: Tang, Maria

; APPLICANT: Loper, John C.
; APPLICANT: Gleeson, Martin
; NAME OF APPLICANT: GLEESON, MARTIN
; ADDRESS: 1000 WASHINGTON ST AND NADEN AVENUE D450

; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
 ; TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
 ; TITLE OF INVENTION: HYDROLYTIC COMPLEX OF CANDIDA TROPICALIS AND METHODS
 ; TITLE OF INVENTION:

TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS
 OF INVENTION: RELATING THERETO
 FILING REFERENCE: 1010-16 sec

```

; FILE REFERENCE: 1010-16.seq
; CURRENT APPLICATION NUMBER: US/09/302,620B
; CURRENT FILING DATE: 1999-04-30

```

```

; CURRENT FILING DATE: 1999-04-30
;
; NUMBER OF SEQ ID NOS: 109
;
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SOFTWARE: PASCALIN VER. 2.1
; SEQ ID NO 99
; LENGTH: 540

```

GEN: 340
TYPE: PRT
ORGANISM: Candida tropicalis

US-09-302-620B-99

Query Match 11.6%; Score 331; DB 4; Length 540;
Best Local Similarity 23.9%; Pred. No. 2.6e-24;

Matches 135; Conservative 94; Mismatches 205; Indels 132; Gaps 22

Qy 42 NHTNYF-----TGFLDILSARTGEE--HAK-----YREKYG 71

Db 25 NYTRWYFIPLVLSLNFISLHTKYLERRFHAKPLGNWLDPPFGIAHPLILYLKSKG 84

Q7	72	STIRFA-----GIAGAPVINSTPKVFNHWKEADYDPKGEWAAR	111
		: : : :	
D5	95	MTAKPACWCEANNVZVTKDQVYVWTCI BUCVI DTETIDIDNTKAVTATGONNDSIGTDRD	144

D6	TVMKFAFSPFNKKYIVRDPKYKTIGLRVGLPFIETIDPENIKAVLATQFNDFSLGRHD
85	
Cv	VPIATGNGIATTAEGEAHKBRRIMIPSI.SAOAVKSMUPTIELEKGMEIVDKMMBDAAEKD
112	
Cv	
171	

QY	112	VRIATGSGVVAEAGEAHRKKRKKIMISLSQAQAVASWVFIFLEAGTETVDTNMTDHRDNR	117
	112		
DB	145	FLYSLRGGIFLDGAGWKSRTMLRPOFAEOWSVH	191
	145		

DD 145 F01SBLTGGDGF 1DDGAGWMSKIMUKFQFAKQVSHV -----NUTSEFHVQVF 133

100

```

; AFFILIANT: Giesch, Nathani
; TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
; TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
; TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS
; TITLE OF INVENTION: RELATING THERETO
; FILE REFERENCE: 1010-16.seq
; CURRENT APPLICATION NUMBER: US/09/302,620B
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 99
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Candida tropicalis
; US-09-302-620B-99

Query Match 11.6%; Score 331; DB 4; Length 540;
Best Local Similarity 23.9%; Pred.No. 2.6e-34;
Matches 135; Conservative 94; Mismatches 205; Indels 132; Gaps 22

CY 42 NHTNYF-----TGNFLDILSARTGEE--HAK-----YREKYG 71

```

25	NYTRWVYIP.LVLLSLNFI.SLLHTKYLREBFHAKPLGNVLDFTFGIATPLILILYLSKG	84
72	STLREA-----GIAGAPVLNSTDPKVFNNHMKKEYADVPKQMAAR	111
85	TYWKFAWGFNNKIVKDPKTKTGLRVLGPLETIDPENIKAVIATQFNDFSLGTRHD	144
112	VLRIATGQGVTAAGEAHKRRHRRIMPSLSAQAVKSMVPIFEKGMELVDKMMEDAAEKD	171
145	FYLSLLGGDIFTLDGAGWGHSRTMLRFPQFAREQVSHV-----KLEPHVQVF	191

QY 172 MAVGESAGEKATLETEGVDVVDKMGVGRATLDVMAAGFDYKSDSLQNTNELYVAVGL 231
Db 192 F-----RHVRKHGQCTDIQELFFRLTVDSATEBFLFGESASLRDSD-----VGL 236
QY 232 TDGFAPLDSFKALMWDVFFVFTMK--RRHEIPTQGLAVSRVVGIELMEQKQOAVLGS 289
Db 237 T-----PITKQEG-RGDFADAFVSYQYQAYRFLQOQWYILN--GAFF--RKSIAVHK 287
QY 290 ASDQAVK-----KDVQGRDILSLVRANIANLPSOKLSDBEVLQ-QISNLLPAGYE 342
Db 288 FADHYQAKALELTDQDQGVVYFEL-----AQOTEDPKVLRDQULNLLVAGRD 339
QY 343 TSSTVLTWMEHRLSEDKAVQDKLEBEI-----CQIDTMDPTLDELNALPYLFAFVKES 395
Db 340 TTAGLLSFVVELSRNEVEFVAKLREVENRFGLGESARVEEISFESLAKSCEYLKAVINEA 399
QY 396 LRLDPPPPYANRECLDKDEDFIPLAEPVIGEDGSVINEVRITKGTWVMLPLFNINRSKFIY 455
Db 400 LRLYPSVPHNFRVATRNIT-LPRGG--GKDG--CSPVIVKGGQVWYVIGTHRDSIY 453
QY 456 GEDAEERPRERWLEDVTDLSNIEAPYGHQASFIISGPRACFGWRFAVAKAFVLTILR 515
Db 454 GADADVFRPRERWPEPTEKLGWAVP-----FNGSPRICLQOQFALTEAS--YVTVRL 504
QY 516 VQ-----FPIISHPEYHITLIS 535
Db 505 LQDFGNLSLOPNAEYPPKQLNTLTL 530

RESULT 10

US-08-457-274A-22
; Sequence 22, Application US/08457274A
; Patent No 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/POCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Musca domestica
; STRAIN: Rutgers

; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: Chromosome 5
US-08-457-274A-22
Query Match 11.68; Score 330.5; DB 1; Length 507;
Best Local Similarity 23.58; Pred. No 2.6e-24;
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;
QY 10 ALGAAPSWASIAFFSYLAPRSSLYNLOGPNH--TNYFTGNFLDILSA-RTGEEHAKY 66
Db 10 ALGV-----LASLALYFVRWNGFYKRGPIHPEEPHLVMGNVKGSLRSKYHIGIADY 62
QY 67 REKYGSTLRFAGI--AGAPVLNSTDPKFNHYV--KEAYDYPKPGMAARVLRATIDGVTY 123
Db 63 YRKPKGSGPAGIFLGHKPAAVLVLDKELKRVLLKDFSNFANRGLYYNEKODPLTGLHVM 122
QY 124 AEGBAHRRHRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMDAABKDMVAGESAGEKA 183
Db 123 VEGEKWSRLTKLSPTFTAGKMKMYNTVLEVGQRLLEVYME----- 164
QY 184 TRLETEG-VDVKQWVGKATLDVMAAGFDYKSDSLQNTNELYV 226
Db 165 -KLEVSSELDMDRLARFNTDVISSVAFGIECNLSRNPEDHFLANGKRSIEVPRHNALIM 223
QY 227 AFVGLTDGFAPLDSFKALMWDVFFVFTMKRHEIPTQGLAVSRVVGIELMEQKQOAV 286
Db 224 AFI-----DSFPEL-----SRKLGMRVLPEDVHQF 248
QY 287 LGSASDAQVD---KKDVQGRDILSLVRANIANLPSOK---LSDBEVLQASINLLFA 339
Db 249 FMSSIKETVDYREKNIRNRDFFDLVLDLK--NNPESISKGLGTENELAAQVFPVFLG 305
QY 340 GYETSSTVLTWMEHRLSEDKAVQDKLEBEI-----CQIDTMDPTLDELNALPYLFAFVKES 395
Db 306 GFTSSSTMGFALYELAQNOQLDLREBVEVDFQFKEDNISYDALMNIPYLDQVLNET 365
QY 396 LRLDPPPPYANRECLDKDEDFIPLAEPVIGEDGSVINEVRITKGTWVMLPLFNINRSKFIY 455
Db 366 LRKYPVGSALTQTLNDY-----VVPENPKYV---LPKGLTVFIPVLGIHYDPDELY 413
QY 456 GEDAEERPRERWLEDVTDLSNIEAPYGHQASFIISGPRACFGWRFAVAKAFVLTILR 515
Db 414 -PNPEEPDFERFSPVWVVKQDSVD-----WLGFGDGPKNICIGRFGKMQSRUGLALVIRH 467
QY 516 VQFE-----PFIISHPE 526
Db 468 FRFTVCSKRTDIPMOINPE 485

RESULT 11

PCT-US95-05758-22
; Sequence 22, Application PC/TUS9505758
; GENERAL INFORMATION:
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05758
; FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Musca domestica
STRAIN: Rutgers
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 5
PCT-US95-05758-22

Query Match 11.6%; Score 330.5; DB 5; Length 507;
Best Local Similarity 23.5%; Pred. No. 2.6e-24;
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;

QY 10 ALGLAASWASIAFSLYAPRSSLNLOGPNH--TNYFTGNFLDILSA-RTGEHAKY 66
Db 10 ALGV-----LASLALYFVRNFGYWKRRGIPHEPFLHVMGVNGLRKYHIGELIADY 62

QY 67 REKYGSTLRFAGI--AGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATGQVVT 123
Db 63 YRKFKGSPFAGIFLGHKPAAVLVDKELRKVLLIKDFSNFANRGLYNEKDDPLTGLHVM 122

QY 124 AGEAHKRRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAVGESAGEKA 163
Db 123 VEGEKWRLTKLSPFTTAGOKMYNTVLEVGORLLEVME-----KINELY 164

QY 184 TRLETEG--VDVKDWGRATLDVMAAGFDYKSDSLQN-----KINELY 226
Db 165 -KLEVSSELDWEDILARENTDVGSAFGIECSLRNPHDRFLAMGRKSIIEVPRHNALIM 223

QY 227 AFVGLTDFGAPTLDSFKAIMDFVPYFRTWKRRHEIPLTQGLAVSRVGVGELMEQKQAV 286
Db 224 API-----DSFPEL-----SRKLGMRVLPEVDHOF 248

QY 287 LGSASDAQVD--KKDVGRODILSILVRANIAANLPESQK-----LSDEEVLQAISNLLFA 339
Db 249 FMSSTKEIVDREKNNIRNDFDLVDLK--NNPESISKLGITFNELAAQVFFVFLG 305

QY 340 GYETSSTVLTWVHRLSDKAVQDLREEI-----CQIDTMDPTLDELNALPYLEAFVKES 395
Db 306 GFETSSTMGFALYELAQOQLRLREEVNEVDQFKEDNISYDALMNPYLDQVLNET 365

QY 396 LRLDPPSPYANRECKDEDFPLAEVPVIGRDSGVINEVRITKGTWVMLPLFNINRSKFIY 455
Db 366 LKRYVGSALTKQTLNDY-----VPHNPKV-----LPKGTILVPIVLGIHYDPELY 413

QY 456 GBDAAEFPRERWLEDVTPSLNSIEAPYGHQASFIISGPRACFGWFAVEMKAFVTLRR 515
Db 414 -PNPEEDFERFSPWVQRDSVD-----WLGFGDGRNRCIGMRFQVQSRGLGLALVIRH 467

QY 516 VQFE-----PIISHPE 526
Db 468 FRFTVCSRTDIPMQINPE 485

RESULT 12
US-09-583-447A-4
; Sequence 4, Application US/09583447A

Patent No. 6645745
GENERAL INFORMATION:
APPLICANT: WOJNOMSKI, Leszek
APPLICANT: GELLNER, Klaus
APPLICANT: EISEL, Regina
TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
FILE REFERENCE: 310115.401
CURRENT APPLICATION NUMBER: US/09/583,447A
CURRENT FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 504
TYPE: PRT
ORGANISM: Homo sapiens
US-09-583-447A-4

Query Match 11.5%; Score 328; DB 4; Length 504;
Best Local Similarity 27.1%; Pred. No. 4.6e-24;
Matches 127; Conservative 75; Mismatches 185; Indels 82; Gaps 18;

QY 68 EKYGSTLRFAGI--AGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATGQVVT 124
Db 66 EKYG---EMWGLYEGQQPMLVIMDPMDIKTVLKECYSVFTNQMLPGLPMGLK-SALSFA 121

QY 125 EGEAHRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAVGESAGEKA 184
Db 122 EDEWKRIRTLSPAFSTVFKFEMVPIISQCQDMLVRSURQAE----- 165

QY 185 RLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDGTFAPTLDSFKA 244
Db 166 --NSKINLKDFGAYTMDVITGLTGVNLDLNNPQD-----PFLKNMKK 209

QY 245 IM-WDFVPVFRWKRRHEIPLTQGLA-VSRRVGIELMEQKQAVLGSASDAQVDK--XDV 300
Db 210 LLKLDLDPPLL-----LSLFPFLTPVFEALNIGLFPKDVTHFLKNSIERKESRLKDK 264

QY 301 QGR--DILSLVRANTAAANLPESQKLSDEEVLQAISNLLFAGYETSSTVLTWVHRLSED 358
Db 265 QKRVDFEQMIQSDNSKETKSHKALSDLELVAQSIIFAAVDTTSTLTPFIMVELATH 324

QY 359 KAVQCKLREICQIDTMDP-----TLDELNALPYLEAFVKESLRDPPSPYANRECLDKE 413
Db 325 PDVQCKLQEE---IDAVLPNKAPVTYDALVQMEYLDVMVNETLRLFPVVSRTVRCKKDI 381

QY 414 DFPLAEVPVIGRDSGVINEVRITKGTWVMLPLFNINRSKFIYGEDAAEFPRP-RWLEDVT 472
Db 382 E-----INGVFTPKGLAVMVPYIALHHD-P-KYWTPEKFCPESRFSKKNK 425

QY 473 DSLNSIE-APYGHQASFIISGPRACFGWFAVEMKAFVTLRRVQFEP 520
Db 426 DSDLYRYIPFG-----AGPRNCIGRFALTNIKLAIVIRALQNFSPKP 468

RESULT 13
US-09-302-620B-98
; Sequence 98, Application US/09302620B
; Patent No. 6331420
GENERAL INFORMATION:
APPLICANT: Wilson, C. Ron
APPLICANT: Craft, David L.
APPLICANT: Erlich, Dudley
APPLICANT: Eshoo, Mark
APPLICANT: Madduri, Krishna M.
APPLICANT: Cornett, Cathy A.
APPLICANT: Brenner, Alfred A.
APPLICANT: Tang, Maria
APPLICANT: Loper, John C.
APPLICANT: Gleeson, Martin
TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
TITLE OF INVENTION: OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS

```
; TITLE OF INVENTION: RELATING THERETO
; FILE REFERENCE: 1010-16 seq
; CURRENT APPLICATION NUMBER: US/09/302,620B
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 98
; LENGTH: 540
; TYPE: PR1
; ORGANISM: Candida tropicalis
US-09-302-620B-98

Query Match      11.5%; Score 328; DB 4; Length 540;
Best Local Similarity 23.9%; Pred. No. 5.1e-24;
Matches 133; Conservative 93; Mismatches 197; Indels 134; Gaps 22;

QY 42 NHTNYF-----TGNFLDILSAATGEE--HAK-----YREKYG 71
DB 25 NTRWYFPLVLLSLNLSLHLYLERRPHAKPLGNFVDPFGIATPLLIIYLSKG 84
QY 72 STLRF-----GIAGAPVLNSTDPKVFNVHMKEAYDYPKPGNAAR 111
DB 85 TWKFWGLWNNKXIVROPKYKTTGLRIVGLPLIETMDPENIKAVLATQNFDSLGTRHD 144
QY 112 VLRIATGCVVTAEGEAKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKD 171
DB 145 FLYSLGDIITDAGWGHSTMLRPOFAREQVSHV-----KLEPHVQVF 191
QY 172 NAVGESAGEKKATRLTEGVVDKDWGRATLDVMAAGFDYKSDSLQNKNELYVAFVGL 231
DB 192 F-----KHVRKRGQTFDIQELFFRLTVDSATEFLFGESASLRDES-----IGL 236
QY 232 TDCGAPLDSFKAIMWDFVYPRTWK--RRHEIPLTQGLAVSRVGTIELMEQKQVAVLGS 269
DB 237 T-----PTTKDFDG--RRDFADAFNYQTYQAVRFLQQMYWILN--GSEF--RKSIAVHK 287
QY 290 ASDQAVDK-----KDVQGRDILSLVZANTAAANLPESQKLSDEVLAA-QISNLLFAGYE 342
DB 288 FADHYVQKALELTDDDLQDQGVVFLYEL-----AKQTRDPKVLRDQLNLVAGRD 339
QY 343 TSTVLTWPHRLSEDKAVQDKLREI-----CQIDTMDPTLDLNAFLYEAFAFKES 395
DB 340 TTAGLSFFYFLSRNPVEFVAKLREVENRFGLEGEARVEISFESKSCYLKAVINET 399
QY 396 LRLDPPSPYANRECKLDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIY 455
DB 400 LRLYSPVPHNFRVATNTT-LRGG-----GEDG--YSPIVVKGVVWYTVIATHRDPISY 453
QY 456 GEDAEFRPRERWLEDVDSLSIEBAPYGHQASFGSPRACFGWRFAVAKMAFLFVTLRR 515
DB 454 GADADVPRPRERWPEPEPRKLGWAYVP-----FNGGPRICLQQQFALTEAS---YVTVRL 504
QY 516 VQPEPIISHPEHITL 532
DB 505 LQ-----EPAHLSM 513

RESULT 14
US-09-158-767-19
; Sequence 19, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; EARLIER FILING DATE: 1998-09-23
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; EARLIER FILING DATE: 1998-09-23
; EARLIER FILING DATE: 1997-12094

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 541
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Altered sequences
US-09-158-767-19

Query Match      11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.6%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

QY 78 GIA--GAPVLNSTDPKVFNVHMKEAYD--YKPGMAARVLRATGCVVTAEGEAKHRR 134
DB 73 GVARGGGLVTVTCPRNLEHLVKARFDNYKPGFWGVRDLGDIFFNSDGDGTWLAQSK 132
QY 135 IMIFSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDQVAVGESAGEKKATRLTEGVVDK 194
DB 133 TAALFTTIRLRTAMSRWVSRIH--GRLLPILA--DAAGKAAQ-----VDLQ 176
QY 195 DWGRATLDVMAAGFDYKSDSLQ--KTNELYVAFVGLTDGPAFTLDSFKAIMWDFVY 252
DB 177 DLLLRLTFDNLICGLAFGKDPETLAQGLPENEFASAFDRATEA---TLNRF-----IFPE 227
QY 253 F-----RTMKERHEIPLTQGLA--VSREVGTELMEOKKQAVLGSASQAVDKDVQGRDIL 306
DB 228 FLWCKKWLGLMGMTTITSSMAHVDQVLAIVIKRKLAAAGKGCDAATHD---DLL 283
QY 307 SLLVRANIAANLPESQKLSDEVLAAQISNLLFAGYETSTVLTWMPHRLSEDKAVODKUR 366
DB 284 SRFMKRG-----SVSDLSLQHVAFILAGRDTSSVALSWPFWLSTHPAVERKIV 334
QY 367 BEICQI-----DTDMPTLDLNAFLYEAFAFKESLRLDPPSPYANRECKLDE 413
DB 335 RELCSVLAASRGADDPALWAEPTFEELRLYLKAASETLRLVSPEDSKHVVD-- 393
QY 414 DFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEPRPERWLE---- 469
DB 394 DYLP-----DOTF-----VPAGSSVTYSIYAGRMKGWGEDCLFEPERWLSADGT 440
QY 470 --DVTDSLNSIEBAPYGHQASFGSPRACFGWRFAVAKMAFLFVTLRRVQFEPILSHPEY 527
DB 441 KFEQHDSYKFF-----AFNAGPRVCLGKLAYLQMKNIAGSVLLRHLRTVAPGRHVE 492
QY 528 EHITLISRPRIVGREKEGYQMRLOVXP 555
DB 493 QKMSLTLEM-----KGG--LRMEVVP 511

RESULT 15
US-09-158-767-20
; Sequence 20, Application US/09158767A
; Patent No. 6180363
; GENERAL INFORMATION:
; APPLICANT: Batard, Yannick
; APPLICANT: Durst, Francis
; APPLICANT: Schalk, Michel
; APPLICANT: Werck-Reichhart, Daniele
; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
; EARLIER FILING DATE: 1998-09-23
; FILE REFERENCE: A32000
; CURRENT APPLICATION NUMBER: US/09/158,767A
; EARLIER FILING DATE: 1998-09-23
; EARLIER FILING DATE: 1997-12094
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 541
; TYPE: PR1
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Altered sequences
US-09-158-767-20

Query Match      11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.6%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

Qy 78 GIA--GAPVLNSTDPKVFNHVKAEYD-YKPCMAARVLRIATGQVVTAEGEAHKRHR 134
Db 73 GVAREGLVTVTCDFRNLHVLKARDNYPKGFHWGVFDLLGDGIFNSDGDWTWLAQEK 132
Qy 135 IMPLSAQAQVKSMPVIFLEKGMELVDKXMDAAEKMDAVGESAGEKKATRLTEGVVK 194
Db 133 TAALFETTRTLRTAMSRWVSRSIH--GRLLPILA--DAAKGRAQ-----VDLQ 176
Qy 195 DWGCRATLDVMALAGFDYKSDSLQN--KTNELVAVFVGLTDGFAPTLDSFKAIMWDFVY 252
Db 177 DLLRLTFDNLGAFGKDPETLAQGLPENEPAFAFDRATEA---TLNRF-----IPPE 227
Qy 253 F-----RTMKRRHEIPLTOGLA-VSRVVGIELMEQKQAVLGSASDAQVDKDKVQGRDIL 306
Db 228 FLWRCKWGLGMEITLTSNAHVQVYLAAVIKRKLELAAGNGKCDTAATHD----DIL 283
Qy 307 SLLVRANIAANLPESQKLSDEVLQAISNLLPAGYETSSVTLTMMFRLSEDKAVQDKLR 366
Db 284 SRFMRKG-----SYDSELOHVALNFILAGROTSSVALSWFFWLSTHPAVERKIV 334
Qy 367 BEICQI-----DTDMPTLDLNLALPYLEAFVKESLRLDPPSPYANRECLKDE 413
Db 335 RELCSVLAASRGADHPALWLAEPFTEELDRVLYKAALSETLRLYPSVPEDSKHVAD- 393
Qy 414 DFIPLAEPVIGRDSVINEVRITGTWMLPLFNINRSKFIYGEDAEERPERWLE----- 469
Db 394 DYLP-----DGTG-----VPAGSSVTYSIYSGRMKGWGEDCLEFRPERWLSADGT 440
Qy 470 -DVTDSLNSIEAPYGHQAFISGFPRACFGWRFAVAKMAFLFTLRRVQFEPFIIISHPY 527
Db 441 KFEQDSYKVV-----AFNAGPRVCLGKDLAYLQMKNIAGSVLLRHRLTVAPGRHVE 492
Qy 528 EHITLIISRPRIVGREKEGYQMRLOVKP 555
Db 493 QKMSLTLEW-----KGG--LRMEVRP 511
```

Search completed: April 2, 2004, 14:02:14
Job time : 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 54 Seconds
(without alignments)
2914.425 Million cell updates/sec

Title: US-10-066-007-1

Perfect score: 2852

Sequence: 1 MFILVLTGALGLAFAWSA.....RIVGREKEGYQMRLOVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2852	100.0	557	3	AAB08713 Amino aci
2	398.5	14.0	503	6	ABU57260 Human chy
3	398.5	14.0	503	6	ABR82024 Human cyt
4	396.5	13.9	502	7	ADD48378 Human Pro
5	396.5	13.9	503	2	AAR72363 Human cyt
6	396.5	13.9	503	2	AAR81464 Human der
7	396.5	13.9	503	2	AAR93170 Human cyt
8	396.5	13.9	503	2	AAY05202 Human CYP
9	394	13.8	503	5	ABG68753 Cytochrom
10	392.5	13.8	527	5	ABBO8079 Maize cyt
11	385.5	13.5	503	5	AAB22852 Human cyt
12	385.5	13.5	535	5	ABG68747 Cytochrom
13	385	13.5	502	5	AAB26192 Human Pro
14	385	13.5	502	7	ADE57186 Human Pro
15	385	13.5	502	7	ADE57190 Human Pro
16	385	13.5	502	7	ADE63919 Human Pro
17	382	13.4	454	5	ABG68754 Cytochrom
18	373.5	13.1	454	6	ADA48312 Rice prot
19	370	13.0	504	7	ADE57188 Rat prote
20	370	13.0	504	7	ADE63917 Rat prote
21	370	13.0	504	7	ADE57184 Rat Prote
22	367	12.9	520	3	AAB23917 Arabidops
23	365	12.8	520	6	ABU54571 Human NOV
24	363.5	12.7	512	5	AAU97096 Abscisic
25	362.5	12.7	465	3	AAG54205 Arabidops

26	362.5	12.7	475	3	AAG54204	AAG54204 Arabidops
27	362.5	12.7	523	3	AAG54203	AAG54203 Arabidops
28	360	12.6	512	2	AAY05898	AAY05898 Vicia sat
29	356.5	12.5	475	3	AAG24685	AAG24685 Arabidops
30	356.5	12.5	522	3	AAG24684	AAG24684 Arabidops
31	356.5	12.5	527	3	AAG24683	AAG24683 Arabidops
32	354	12.4	562	4	AAU30319	AAU30319 Novel hum
33	351	12.3	512	5	AAU97097	AAU97097 Abscisic
34	350.5	12.3	1115	6	AAO16061	AAO16061 Gibberell
35	350.5	12.3	1054	6	AAO16055	AAO16055 Bacillus
36	350.5	12.3	1054	6	AAO16056	AAO16056 Bacillus
37	349	12.2	508	5	ABB07528	ABB07528 Human dru
38	348.5	12.2	511	7	ADE60554	ADE60554 Human Pro
39	348.5	12.2	546	4	AAU14298	AAU14298 Human nov
40	347.5	12.2	524	4	AAU14534	AAU14534 Human nov
41	347.5	12.2	524	4	AAU14536	AAU14536 Human nov
42	347.5	12.2	524	4	AAU14535	AAU14535 Human nov
43	346.5	12.1	511	7	ADE60552	ADE60552 Rat Prote
44	346.5	12.1	524	3	RAY65706	RAY65706 Membrane-
45	346.5	12.1	524	4	AAB73677	AAB73677 Human Oxi

ALIGNMENTS

RESULT 1

AAB08713
ID AAB08713 standard; protein; 557 AA.

XX
AC AAB08713;

XX
AC AAB08713;

DT 12-SEP-2003 (revised)

DT 02-JAN-2001 (first entry)

XX
AC AAB08713;

DE Amino acid sequence of an astaxanthin synthetase polypeptide.

XX Astaxanthin synthetase; astaxanthin; beta-carotene; carotenogenic yeast;

KW Antioxidant; cancer; colouring reagent; farmed fish; salmon.

XX Xanthophyllomyces dendrorhous.

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

XX
AC AAB08713;

Query Match		100.0%; Score 2852; DB 3; Length 557;
Best Local Similarity		100.0%; Pred. No. 4.2e-240;
Matches 557; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MFILVLLTGALGLAFAFWASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDLSARTG 60
DB	1	MFILVLLTGALGLAFAFWASIAFFSLYLAPRRSSLYNLQGNHNTYFTGNFLDLSARTG 60
QY	61	EEHAKYREKYGSTLRPAGIAGAPVLNSTDPKVFNHMKVKEAYDYPKPGVAARVLIATGG 120
DB	61	EEHAKYREKYGSTLRPAGIAGAPVLNSTDPKVFNHMKVKEAYDYPKPGVAARVLIATGG 120
QY	121	VVTABEAEHKKRRIRIMIPSLSAQAVKSMVIPFLFKGMELVDKMMEDAAEKDMVAGESAGE 180
DB	121	VVTABEAEHKKRRIRIMIPSLSAQAVKSMVIPFLFKGMELVDKMMEDAAEKDMVAGESAGE 180
QY	181	KKATRLTEGVDKWDVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDFAPTLTLD 240
DB	181	KKATRLTEGVDKWDVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDFAPTLTLD 240
QY	241	SPKAIMWDVFPVPRTRMKRHEIPLTQGLAVSRVGIEMKQKQAVLGSDQAVDKDV 300
DB	241	SPKAIMWDVFPVPRTRMKRHEIPLTQGLAVSRVGIEMKQKQAVLGSDQAVDKDV 300
QY	301	QGRDILSLVVRANIAANLPESQKLSDEEVLQISNLLFAGYETSTVLTWPHRLSBDKA 360
DB	301	QGRDILSLVVRANIAANLPESQKLSDEEVLQISNLLFAGYETSTVLTWPHRLSBDKA 360
QY	361	VQKLEEEICQIDTMDPTLDENALPYLEAFVKESRLDPPSPVANRECLDEDFIPLAE 420
DB	361	VQKLEEEICQIDTMDPTLDENALPYLEAFVKESRLDPPSPVANRECLDEDFIPLAE 420
QY	421	PVIGRSGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERFERWLEDVTDLSNIEA 480
DB	421	PVIGRSGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERFERWLEDVTDLSNIEA 480
QY	481	PYGHQASFTSGPRACGFWFAVAKELFVTLRRVQPEPIISHPEYEHITLIISRPIV 540
DB	481	PYGHQASFTSGPRACGFWFAVAKELFVTLRRVQPEPIISHPEYEHITLIISRPIV 540
QY	541	GREKEGYQMLQVKPVE 557
DB	541	GREKEGYQMLQVKPVE 557
RESULT 2		
ABUS7260		
ID	ABUS7260	standard; protein; 503 AA.
XX	XX	ABUS7260;
AC	AC	ABUS7260;
XX	XX	25-APR-2003 (first entry)
DT	DT	Human chytochrome P450 CYP3A4 protein.
DE	DE	
XX	XX	Human; enzyme; transgenic; drug metabolism; behaviour; mouse;
KW	KW	pharmacokinetic assay; pharmacodynamic assay; toxicology; serum albumin;
KW	KW	alpha-acidic glycoprotein; CYP; multidrug resistance protein; MRP;
KW	KW	uridine diphosphoglucuronosyl transferase, UGT; cytochrome P450.
XX	XX	
OS	OS	Homo sapiens.
XX	XX	
FN	FN	WO200283897-A1.
XX	XX	24-OCT-2002.
PD	PD	
XX	XX	18-APR-2002; 2002WO-AU000485.
EF	EF	
XX	XX	18-APR-2001; 2001AU-00004467.
PR	PR	
XX	XX	(GENE-) GENE STREAM PTY LTD.
PA	PA	
XX	XX	

PI	Daly JV;	
XX	XX	
DR	WPI: 2003-093021/08.	
DR	N-PSDB; ABX77172.	
XX	XX	
PT	New transgenic non-human animal expressing a foreign polypeptide associated with drug behavior and/or metabolism, useful for studying the behavior and/or metabolism of a drug in other animals.	
PT	PT	
XX	XX	
PS	Disclosure; Page 299-301; 408pp; English.	
XX	XX	
CC	This invention relates to a transgenic non-human animal which may be used for assessing the behaviour and/or metabolism of a drug in another animal and which expresses a foreign polypeptide associated with drug behaviour and/or metabolism. The invention also comprises a nucleic acid construct for use in producing the above transgenic non-human animal and a method of assessing the metabolism and/or behavior of a drug in an animal of interest, comprising administering a test agent to the transgenic animal and conducting analytical tests to determine drug metabolism and/or behavior. The transgenic animal is useful in studying drug metabolism and/or behavior in other animals. The nucleic acid construct is useful in producing the above transgenic animal and the methods are used for producing, breeding and using transgenic animals for pharmacological (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological studies. Nucleic acid sequences used within the invention are serum albumin; alpha-acidic glycoprotein; cytochrome p450 (CYP); uridine diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins and (MRP's). The present sequence represents a protein sequence used to create a transgenic animal within the scope of the invention	
CC	CC	
XX	XX	
SQ	Sequence 503 AA;	
Query Match		14.0%; Score 398.5; DB 6; Length 503;
Best Local Similarity		28.7%; Pred. No. 8.6e-26;
Matches 153; Conservative		88; Mismatches 195; Indels 97; Gaps 24;
QY	13	LAAFSW--ASIAFTSLYLAPRRS--SLYNLQO--PNHTNY--FTGNFLDLSARTG-----E 61
DB	7	LAMETWLLAVSLVLLLYLGYTHSHGLFKLGIQPTPLPLGN---ILSYHKGFCMFDM 63
QY	62	EHAKEYREKYGSTLRPAGIAGAPVLNSTDPKVFNHV--KEAYDY---PKGMAARVLIAT 117
DB	64	CHKYKVMGF--YDG--QQPVLAITDDPMIKTVLVKECYSVFTNRRPPGPGVFKMSA- 117
QY	118	GDGVVTAEGEAHKKRRIRIMIPSLSAQAVKSMVIPFLFKGMELVDKMMEDAAEKDMVAGES 177
DB	118	---ISIAEDEBKRLSLLSPTFTSGKLEWVPIIAQYGDVLYNLRREA----- 164
QY	178	AGEKATLET--EGVDVKDWVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTDFGA 236
DB	165	-----ETGKPVTLKDVFGAYSDVITSTSGVNIIDSLNNPD----- 201
QY	237	PTLDSFKAIM--WDFV--PYFRTMK--RRHEIPLTQGLAV---SRRVGIELMEQKQAVLGSA 290
DB	202	PFVENTKLLRFDLDPFFLSITVPFPLILEVLNLCVFPREVNTFLRSVXR-----M 256
QY	291	SDQAYDKDVGQRDILSLVVRANIAANLPESQKLSDEEVLQISNLLFAGYETSSVLTW 350
DB	257	KESRLEDTQKRVDFQLQMLDMSQSKETESHKALSDLELVAQSIIFIFAGYETSSVLSF 316
QY	351	MFHRLSEDKAVQDKLREIICOI--DTDMPTLDENALPYLEAFVKESRLDPPSPVANRE 408
DB	317	IMYELATHPDVQOKLQEEIDAVLPKAPPTDYDTVLQMEYLDVWVNEITLRFPLAMLERV 376
QY	409	CLKDEDFPLAEPVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEERFERWL 468
DB	377	CKKQVE-----INGMFIPKGVVWVMPYSALHRDP--KYWTBPEKFLPERFS 420
QY	469	EDVTDLSNS--IEAPVGHQASFTSGPRACGFWFAVAKMAFLFVTLRRVQPEP 520
DB	421	KKNKONIDFYITPFG-----SGFRNCIGRFAAMNKKLALIRVLQNFSEK 467

```

RESULT 3
ABR82024
ID ABR82024 standard; protein; 503 AA.
XX
XX ABR82024;
XX
DT 22-SEP-2003 (first entry)
XX
XX Human cytochrome P450 3A4 amino acid sequence.
XX
XX Human; protein array; protein moiety; phenotype; drug discovery;
XX naturally occurring variant; pharmacogenomic; diagnostic;
XX parallel analysis; tumour suppressor; p53; cytochrome P450.
XX
XX OS Homo sapiens.
XX
XX WO2003048768-A2.
XX
XX 12-JUN-2003.
XX
XX 05-DEC-2002; 2002WO-GB005499.
XX
XX 05-DEC-2001; 2001US-0335806P.
XX
XX 16-SEP-2002; 2002US-0410815P.
XX
XX (SENS-) SENSE PROTEOMIC LTD.
XX
XX Boutell JM, Godber BLJ, Hart DJ, Blackburn JD;
XX
XX WPI; 2003-569063/53.
XX
XX N-PSDB; ACF06054.
XX
XX
XX New protein array, useful for determining the phenotype of a naturally
XX occurring variant of a DNA sequence of interest, comprises a surface upon
XX which at least two protein moieties are deposited.
XX
XX Example 5; Fig 11B; 84pp; English.
XX
XX The present invention describes a protein array comprising a surface upon
XX which at least two protein moieties are deposited at spatially defined
XX locations, where the protein moieties are naturally occurring variants of
XX a DNA sequence of interest. Also described: (1) making a protein array;
XX (2) screening a set of protein moieties for molecules that interact with
XX one or more proteins; and (3) simultaneously determining the relative
XX properties of members of a set of protein moieties. The protein array can
XX be used for determining the phenotype of a naturally occurring variant of
XX a DNA sequence of interest. The protein array is useful for drug
XX discovery, pharmacogenomics and diagnostics. The protein array allows the
XX parallel analysis of closely related proteins with a sensitivity that is
XX at least comparable to existing methods, if not better, with small
XX volumes of potentially expensive ligands, and in a quantitative,
XX comparative functional analysis manner not previously possible. ACF06000
XX to ACF06056 and ABR81975 to ABR82026 represent sequences used in the
XX exemplification of the present invention
XX
XX Sequence 503 AA;
XX
XX Query Match 14.0%; Score 398.5; DB 6; Length 503;
XX Best Local Similarity 28.7%; Pred. No. 8.6e-26;
XX Matches 153; Conservative 86; Mismatches 195; Indels 97; Gaps 24;
XX
XX 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLOG--PNHTNY--FTGNFLDILSARTG-----E 61
XX 7 LAMETWLLAVSLVLLYLGTHSHGLKGLGIPGPTPLPLGN---ILSVHKGFCMFDM 63
XX
XX 62 EHAKVREKYGSTLRFAGIAGPVLNSTDPKVNHVH--KEAYDY---PKGMAARVLRAT 117
XX 64 CHKXIKGVWGF---YDG--QQPVLAITDPMIKTVLKECYSVFTNRRFGPVGFKMSA- 117
XX
XX 118 GDGVVTVAGEAHRHRIMIPSLPSAQAVKSMVPFIFLEKMGELVDYDKMEDAEKDMVGES 177
XX 118 ---TSIAEDBEKWLRLSILSTFTSGKLKEMWPIAQGVGVLRNLRRA----- 164

```

CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 502 AA;

Query Match 13.9%; Score 396.5; DB 7; Length 502;
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
 Db 6 LAMETWLLAVSLVLLYGLTHSHGLPKLGPPTPLPFLGN---ILSYHKGFCMPDME 62
 QY 62 EHAQYREKYGSTLRAGAPVLNSTDPKFNHVM--KEADY---PKPGMAARVLRIAT 117
 Db 63 CHKYGKVGWF---YDG--QQPVLAITDPMIKTVLVEKCVFTNRRPFGVGFMSA- 116
 QY 118 GDGVVTAEGEAHRRHRRIMIPSLSAQAVKSMVPIFLEKGMELVDQMDEAAEKDMAYGES 177
 Db 117 ---ISIAEDEEMKRLSLLSPTFTSGKLEWVPITIAQYGDVLRNLRREA----- 163
 QY 178 AGEKKATRIET-EGVDVKQWVGTRATLDVVALAGPDYKSDSLQNKTNELYVAFVGLTDGFA 236
 Db 164 -----ETGKPVTLKDVFGAYSDMTVITSTSGVINDSLNNPD----- 200
 QY 237 PTLDSFKAIM-WDFV-PYFRTWK-RRHEIFLTQGLAV---SRRVGIELMEQKQAVLGSA 290
 Db 201 PFVENTKLLRFDLDFDPLFISITVFPFLIPILEVNICVFPREVTNLRKSVK- ---M 255
 QY 291 SDQAVDKQVQGRDILSLVLRANIANLPESQKLSDEEVLQAISNLLFAGYETSSVLTW 350
 Db 256 KESRLDQTKHRVDFLQMLDMSQSKETESHKALSDLELVAQSIIFIFAGYETSSVLSF 315
 QY 351 MFHRLSDEKAVQKLRREEICQI--DTPMTLDLNLALPYLEAFVKESLRLLDPPSPYANRE 408
 Db 316 IMVELATHPDVQKQLEEDDAVLPNKAPPTVTLQMEYLDWMVNETLRPIAMRLERV 375
 QY 409 CLKDEDFIPLAEPVIGRDSGVINEVRIITGTWVMPLFNINRSKIYGEDAEERFPERWL 468
 Db 376 CKKDVE-----INGMFIKGVVMVPIPSYALHRDP-KYWTPEKEFLPERFS 419
 QY 469 EDVTDLSNS-IEAPYGHQAFISGPRACFGWFAVAEMKAFVTLRLRVQEPF 520
 Db 420 KKKONIDPIYIYTFPG-----SGPNCIGMRFALMMKALIRVLQNFSEKFP 466

RESULT 5
 AAR72363
 ID AAR72363 standard; protein; 503 AA.
 XX
 AC AAR72363;
 XX
 DT 25-MAR-2003 (revised)
 DT 14-NOV-1995 (first entry)
 XX
 DE Human cytochrome P450 molecular species 3A4 protein.
 XX

KW Human cytochrome P450; amplification; PCR; primer; expression vector;
 KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
 KW carcinogen; mutagen; liver metabolism.
 XX Homo sapiens.
 PN EP644267-A2.
 XX
 PD 22-MAR-1995.
 XX
 PF 20-JUL-1994; 94EP-00111298.
 XX
 PR 20-JUL-1993; 93JP-00201120.
 PR 21-JUL-1993; 93JP-00180246.
 PR 30-JUL-1993; 93JP-00208279.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 PI Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
 FI WPI; 1995-116991/16.
 DR N-PSDB; AAQ87717.
 XX
 PT Evaluation of safety of a chemical cpd. - using recombinant yeast
 PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
 XX
 PS Example; Page 31-33; 124pp; English.
 XX
 CC The amino acid sequence of the human cytochrome P450 species 3A4. The
 CC cDNA was amplified by PCR using the primers AAQ87743-6. The product was
 CC cloned into the yeast expression vectors pAAH5N or pAHR5 to produce the
 CC vectors p3A4 for the expression of the cytochrome P450 alone or p3A4R for
 CC co-expression with the yeast NADPH-P450 reductase. The vectors are used
 CC in a method for evaluating the safety of a chemical compound by reacting
 CC the chemical compound with recombinantly produced human cytochrome P450
 CC molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4,
 CC or their auxiliary species and variants (AAQ87718-32), and yeast NADPH-
 CC P450 reductase, either as a fused protein or in cell extracts, and
 CC analysing the resulting metabolite to assess the safety of the chemical
 CC compound. The method is useful for determining whether the chemical
 CC compound, or its metabolite, will be converted into a carcinogenic or
 CC mutagenic form through metabolism in the liver. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX
 SQ Sequence 503 AA;
 Query Match 13.9%; Score 396.5; DB 2; Length 503;
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
 Db 7 LAMETWLLAVSLVLLYGLTHSHGLPKLGPPTPLPFLGN---ILSYHKGFCMPDME 63
 QY 62 EHAQYREKYGSTLRAGAPVLNSTDPKFNHVM--KEADY---PKPGMAARVLRIAT 117
 Db 64 CHKYGKVGWF---YDG--QQPVLAITDPMIKTVLVEKCVFTNRRPFGVGFMSA- 117
 QY 118 GDGVVTAEGEAHRRHRRIMIPSLSAQAVKSMVPIFLEKGMELVDQMDEAAEKDMAYGES 177
 Db 118 ---ISIAEDEEMKRLSLLSPTFTSGKLEWVPITIAQYGDVLRNLRREA----- 164
 QY 178 AGEKKATRIET-EGVDVKQWVGTRATLDVVALAGPDYKSDSLQNKTNELYVAFVGLTDGFA 236
 Db 165 -----ETGKPVTLKDVFGAYSDMTVITSTSGVINDSLNNPD----- 201
 QY 237 PTLDSFKAIM-WDFV-PYFRTWK-RRHEIFLTQGLAV---SRRVGIELMEQKQAVLGSA 290
 Db 202 PFVENTKLLRFDLDFDPLFISITVFPFLIPILEVNICVFPREVTNLRKSVK- ---M 256
 QY 291 SDQAVDKQVQGRDILSLVLRANIANLPESQKLSDEEVLQAISNLLFAGYETSSVLTW 350
 Db 257 KESRLDQTKHRVDFLQMLDMSQSKETESHKALSDLELVAQSIIFIFAGYETSSVLSF 316

QY 351 MFHRLSEDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAFVKSLELDPPSPYANRE 408
 Db 317 IMYELATHPDVQCKLOEEDAVLPNKAPPTDYTLQMEYLDVWVNETLRLFFIAMRLERV 376
 QY 409 CLKDEDFIPLAEPVIGRDGVSINEVRITKGTVMVLPFLFNINRSKFIYGDABEERPERWL 468
 Db 377 CKKQVE-----INGMFIPKGVWVMIPISYALHRDP-KYWTEPEKFLPERFS 420
 QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACGWRFAVAEMKAFLEVTLLRRVQPEP 520
 Db 421 KKKNDNDPIYTPFG-----SGPRNCIGRFPALMMKALIRVLQNFSEKP 467

RESULT 6
 AAR81464
 ID AAR81464 standard; protein; 503 AA.
 AC AAR81464;
 XX
 DT 01-AUG-1996 (first entry)
 XX
 DE Human derived cytochrome P450A4.
 XX
 KW Human derived cytochrome; P450A4; commercial cDNA library; yeast;
 KW transfection; recombinant production; expression vector; mammal;
 KW immunisation; sensitisation; antibody; determination; detection;
 KW non-cross reactive.
 XX
 OS Homo sapiens.
 XX
 FN JP08027197-A.
 XX
 PD 30-JAN-1996.
 XX
 PF 13-JUL-1994; 94JP-00161552.
 XX
 PR 13-JUL-1994; 94JP-00161552.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 DR WPI; 1996-136338/14.
 DR N-PSDB; AAT17399.
 XX
 PT Antibody recognising human derived cytochrome P450A4 - allows specific
 PT detection of cytochrome P450 species in humans.
 XX
 PS Example 1; Page 10-12; 13pp; Japanese.
 XX
 CC The present sequence is the human derived cytochrome (HDC) P450A4, which
 CC was obtd. from a commercial cDNA library. Yeast were transfected with an
 CC expression vector contg. the HDC cDNA, cultured and then disrupted to
 CC give a microsomal fraction. The HDC was purified from the fraction, and
 CC used to immunise and sensitise a mammal. Blood was drawn from the mammal,
 CC and an anti-HDC antibody isolated. The antibody obtd. recognises HDC
 CC P450A4, partic. at a serum dilution rate of 1:10000, and is
 CC substantially without cross reaction to other HDC P450 spp
 XX
 SQ Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAATSW--ASTAFSLYLAPERS-SLYNLOG-PNHTNY-FTGNFLDILGARTG-----E 61
 Db 7 LAMETWLLAVSLVLLLYGTHSHGLFKKIGIGPOTPLPLGN--ILSVHKGFCNFDNE 63
 QY 62 EBAKYREKYGSTLRPAGIAGAPVLNSTDPKVFNVHM-KEAYDY---PKPGMAARVLRIAT 117
 Db 64 CHKXKYGKVGWGF---YDGG--QQPVLAITDPDMIKTVLVKCYSVFTNRPPFGPVGMKSA- 117
 QY 118 GDGVVTAEGEAKHRRIMIFSLSAQAQVSKVMPFLEKGMELVDKWMEDAAEKDMAVGES 177

Db 118 ---ISTAEDENKLSLSLSPFTTSGKLKEMVPIIAQYGDVLRNLRREA----- 164
 QY 178 AGEKKATRLT-EGVDVKDWGRALDYMALAGPDYKSDSLONKTNELYVAFVGLTDGFA 236
 Db 165 -----ETQKPTLKDVFAGYSMDVITSTSGVNIIDSLNPOD----- 201
 QY 237 PTLDSFKAIM-WDFV-PVFRITMK-RRHEIPTLQGLAV---SRRVGIELMEQKQAVLGSA 290
 Db 202 PFVNTKLLARDFDLDPPLSLITVPFLIPILEVLNICVFPREVINFURKSVKR- -M 256
 QY 291 SDQAVKDKVQGRDILSLVRANIANLPESOKLSDEVLQAIQSNLLPAGYETSTVLTW 350
 Db 257 KESRLDTQKRVDFLQMLIDSONSKESHKALSDLEVAQSIFIFAGYETSTSSLSF 316
 QY 351 MFHRLSEDKAVQDKLREICQI--DTDMPTLDLNLALPYLEAFVKSLELDPPSPYANRE 408
 Db 317 IMYELATHPDVQCKLOEEDAVLPNKAPPTDYTLQMEYLDVWVNETLRLFFIAMRLERV 376
 QY 409 CLKDEDFIPLAEPVIGRDGVSINEVRITKGTVMVLPFLFNINRSKFIYGDABEERPERWL 468
 Db 377 CKKQVE-----INGMFIPKGVWVMIPISYALHRDP-KYWTEPEKFLPERFS 420
 QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACGWRFAVAEMKAFLEVTLLRRVQPEP 520
 Db 421 KKKNDNDPIYTPFG-----SGPRNCIGRFPALMMKALIRVLQNFSEKP 467

RESULT 7
 AAR93170
 ID AAR93170 standard; protein; 503 AA.
 AC AAR93170;
 XX
 DT 11-OCT-1996 (first entry)
 XX
 DE Human cytochrome P450 molecular species 3A4 protein.
 XX
 KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
 KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
 KW evaluation; safety; fusion protein; metabolite; detoxification;
 KW carcinogenic.
 XX
 OS Homo sapiens.
 XX
 FN JP08056635-A.
 XX
 PD 05-MAR-1996.
 XX
 PF 15-JUL-1994; 94JP-00164184.
 XX
 PR 20-JUL-1993; 93JP-00201120.
 PR 30-JUL-1993; 93JP-00208279.
 PR 17-JUN-1994; 94JP-00136053.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 DR WPI; 1996-182311/19.
 DR N-PSDB; AAT28383.
 XX
 PT Novel method for the evaluation of the safety of a cpd. - using a human
 PT cytochrome P450 and yeast NADPH reductase to determine whether the
 PT analyte cpd. is detoxified or metabolised to a carcinogen.
 XX
 PS Example 1; Page 24-26; 74pp; Japanese.
 CC This is the amino acid sequence of the human cytochrome P450 molecular
 CC species 3A4 protein. The corresp. gene was amplified from a human liver
 CC derived cDNA library as 2 fragments of 0.6 and 0.9 kb using primers
 CC AAT26933-6. The prod. was cloned into the yeast expression vector pAAV5N
 CC to generate plasmid p3A4 for prodn. of the cytochrome only or into the
 CC vector pAHRR to generate the plasmid p3A4R for co-prodn. with the yeast
 CC NADPH-P450 reductase. The sequence is placed under control of the yeast

CC ADH gene promoter and terminator. The vectors are used in a method for
CC evaluating the safety of a cpd. by reacting the test cpd. with
CC recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380),
CC 2C9 (AAT28381), 2E1 (AAT28382), 3A4 or their variants (AAT28384-98)
CC together with yeast NADPH-P450 reductase (either as a fused protein or as
CC a cell extract) and analysing the resulting metabolite. The cpd. is
CC considered "safe" if it is detoxified or not rendered carcinogenic or
CC "unsafe" if it is not detoxified or is metabolised to a carcinogenic cpd
XX
SQ Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;
Best Local Similarity 28.7%; Pred. No. 1.3e-25;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQG--PNHTNY--FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLAVSLVLLYLGTHSHGLFKLG:PGTLPPLFNG--ILSYHKGFCMFDM 63
QY 62 EHAKYREKYGSLRFRAGIAGAVLNSTDPKVFNVHM--KEAYDY--PQGMARVLRAT 117
DB 64 CHKKYKVGWGF--YDG--QQPVLATDPDMIKTVLVEKCYSVFTNRRPFGPVGFMKSA- 117
QY 118 GDGVVTAEGEAKHRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAGES 177
DB 118 ---ISIAEDEEWKELRSLSPFTTSGKLEWVPIIAQYGDVLRNLRREA----- 164
QY 178 AGEKKAATLET--EGVDVKWVGRATLDVWALAGFDYKSDSLQNKTNELYVAFVGLTDGFA 236
DB 165 -----ETGKPVTLKDVFGAYSDMTSTSGVNISSLNPNQD----- 201
QY 237 PTLDSFKAIM-WDFV-PYFRMTK--RRHEIPLTQGLAV---SRRVGIELMEQKQAVLGSA 290
DB 202 PFVENTKLLRFDLDPFLSITVFPFLIPILEVLNVCVPPREVTNFRKSVKR-----M 256
QY 291 SDQAVDKDVQGRDILSLVRANIANLPESQKLSDEEVLAQISNLLFAGYETSSTVLTW 350
DB 257 KESRLEDTQKRVDFLQIMDSQNSKETESHKALSDELVAQSIIFIFAGYETTSSVLSF 316
QY 351 MFHRLSEDKAVQDKLREEICQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408
DB 317 IMYELATHPDVQKQLEEDAVLPNKAPPTVDVLQWEYLDVWVNETLRFPIAMLERV 376
QY 409 CLKDEDFIPLAEPIVGRDGSVINEVRITKGTMMWMLPLFNINRSKFYIGEDAEFRPERWL 468
DB 377 CKKQVE-----INGMFIPKGVWVMPISYALHRDP--KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRRVQFEP 520
DB 421 KKNKNIDPIYITPFG-----SGPRNCIGMRFALNMNKLALIRVLQNFSPK 467

RESULT 8
AAV05202
ID AAV05202 standard; protein; 503 AA.

XX AC AAV05202;
XX DT 17-JUN-1999 (first entry)
XX DE Human CYP3A4 protein.
XX KW CYP3A4 gene polymorphism; polymorphic locus; human; altered metabolism;
XX KW CYP3A4 substrate; drug-drug interaction identification; toxin exposure;
XX KW Genetic linkage detection; phenotypic variation.
XX OS Homo sapiens.
XX PN WO9913106-A1.
XX PD 18-MAR-1999.
XX PF 02-SEP-1998; 98WO-US018158.

XX 10-SEP-1997; 97US-0058612P.
XX (AXYS-) AXYS PHARM INC.
XX PI Lichter JB, Guida M;
XX WPI; 1999-215070/18.
XX DR N-PSDB; AAX28295.
XX
XX New isolated CYP3A4 polymorphic sequences.
XX
XX Disclosure; Page 25-27; 40pp; English.
XX
XX This sequence represents the human CYP3A4 gene protein. The invention
XX relates to a CYP3A4 sequence polymorphism, which is part of a non-
XX naturally occurring chromosome. Nucleic acids comprising the CYP3A4
XX polymorphic sequences can be used to screen patients for altered
XX metabolism for CYP3A4 substrates, potential drug-drug interactions, and
XX adverse/side effects as well as diseases that result from environmental
XX or occupational exposure to toxins. They can also be used to establish
XX animal, cell culture and in vitro cell-free models for drug metabolism.
XX Polymorphic CYP3A4 gene sequences can be used for expression studies to
XX determine the effect of promoter and/or intron sequence variations on
XX mRNA expression and stability. The polymorphisms are also used as single
XX nucleotide polymorphisms to detect genetic linkage to phenotypic
XX variation in activity and expression of CYP3A4. The nucleic acids can
XX also be used to generate genetically modified non-human animals or site
XX specific gene modifications in cell lines
XX
XX Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;
Best Local Similarity 28.7%; Pred. No. 1.3e-25;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
QY 13 LAAFSW--ASIAFFSLYLAPRRS--SLYNLQG--PNHTNY--FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLAVSLVLLYLGTHSHGLFKLG:PGTLPPLFNG--ILSYHKGFCMFDM 63
QY 62 EHAKYREKYGSLRFRAGIAGAVLNSTDPKVFNVHM--KEAYDY--PQGMARVLRAT 117
DB 64 CHKKYKVGWGF--YDG--QQPVLATDPDMIKTVLVEKCYSVFTNRRPFGPVGFMKSA- 117
QY 118 GDGVVTAEGEAKHRRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMWEDAAEKDMVAGES 177
DB 118 ---ISIAEDEEWKELRSLSPFTTSGKLEWVPIIAQYGDVLRNLRREA----- 164
QY 178 AGEKKAATLET--EGVDVKWVGRATLDVWALAGFDYKSDSLQNKTNELYVAFVGLTDGFA 236
DB 165 -----ETGKPVTLKDVFGAYSDMTSTSGVNISSLNPNQD----- 201
QY 237 PTLDSFKAIM-WDFV-PYFRMTK--RRHEIPLTQGLAV---SRRVGIELMEQKQAVLGSA 290
DB 202 PFVENTKLLRFDLDPFLSITVFPFLIPILEVLNVCVPPREVTNFRKSVKR-----M 256
QY 291 SDQAVDKDVQGRDILSLVRANIANLPESQKLSDEEVLAQISNLLFAGYETSSTVLTW 350
DB 257 KESRLEDTQKRVDFLQIMDSQNSKETESHKALSDELVAQSIIFIFAGYETTSSVLSF 316
QY 351 MFHRLSEDKAVQDKLREEICQI--DTDMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408
DB 317 IMYELATHPDVQKQLEEDAVLPNKAPPTVDVLQWEYLDVWVNETLRFPIAMLERV 376
QY 409 CLKDEDFIPLAEPIVGRDGSVINEVRITKGTMMWMLPLFNINRSKFYIGEDAEFRPERWL 468
DB 377 CKKQVE-----INGMFIPKGVWVMPISYALHRDP--KYWTEPEKFLPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFTSGPRACFGWRFAVAEMKAFVTLRRVQFEP 520
DB 421 KKNKNIDPIYITPFG-----SGPRNCIGMRFALNMNKLALIRVLQNFSPK 467

RESULT 9
ABG68753
ID ABG68753 standard; protein; 503 AA.
XX
AC ABG68753;
XX
DT 07-OCT-2002 (first entry)
XX
DE Cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 2-13.
XX
KW Cytochrome P450; CYP3A1; CYP3A2; CYP3A43; CYP3A4; CYP3A5; CYP3A7;
XX drug metabolism; drug design; drug screening.
XX
OS Homo sapiens.
XX
FN WO200244213-A1.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-SE002631.
XX
PR 28-NOV-2000; 2000SE-00004366.
XX
PR 11-JUN-2001; 2001SE-00002061.
XX
XX (ZAPH/) ZAPHIROPOULOS P G.
XX (FINT/) FINTA C.
XX
XX Zaphiropoulos PG, Finta C;
XX
XX WPI; 2002-557532/59.
XX N-PSDB; ABK97692.
XX
XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
XX of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
XX metabolism, in drug design and drug screening.
XX
XX Claim 2; Fig 11; 131pp; English.
XX
XX The invention describes a cytochrome P450 protein (I) in which CYP3A43
XX exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
XX fragments, variants and multiples of (I) having essentially the same
XX characteristics. (I) is useful as a medicament, and for evaluating drug
XX metabolism, in drug design, and drug screening, and in tests for
XX adjusting the dose of drugs. This is the amino acid sequence of a novel
XX cytochrome P450 protein
XX
SQ Sequence 503 AA;

Query Match 13.8%; Score 394; DB 5; Length 503;
Best Local Similarity 28.7%; Pred. No. 2.1e-25;
Matches 142; Conservative 83; Mismatches 177; Indels 92; Gaps 20;

Qy 47 FTGNFLDILSARTG-----EEHAKYREKYGSTLRPAGIAPVLNSTDPKVFNVHM-KE 99
Db 46 FLGN---ILSYHKGFCMFDMCKYKGVKMGF---YDG--QQFVLAITDPDMIKTVLVE 97
Qy 100 AYDY---PKGMAARVLRIATGCVVTAGEAKHRRIMIPSLSAQAQVKSVPFLFLEKG 156
Db 98 CYSVFTNRRPFGVGFNKA-----ISIADEEHWRLSLSTFTSGKLKEMVPIAQYG 153
Qy 157 MELVDKMDAABKNAVGSAGEKATRLT-EGVDVVDWVGRTLDVMAAGFDYKSD 215
Db 154 DVLVRLNLRREA-----ETGKPVTLKDVFGAYSMVDITSTSGVNID 194
Qy 216 SLQKNTNELVAFVGLTDGAPFLDSFKAIM-WDFV-PYFRIMK-RRHEIPLTQGLAV-- 270
Db 195 SLNPNQD-----PVENTKLLRDFLDFPFLFTVFPFPIPLEVINICV 240
Qy 271 -SRVGIEMEQKQAVLGSASQAVDKQVQGRDILSLVRANIAANLPESQKLSDEEV 329
Db 241 FPREVTFLEKSVKR-----MKESRLDITQKRVDFLQMLDMSQNSKETESHKALSLEL 295
Qy 330 LAQISNLLFAGYETSSVLTWTFMRHUSEKXAVQDKUREBICQI--DTDMPTDELNALPY 387

Db 296 VAQSIIFAGYETSSVLSFIMYELATHPDVQKQLEIDAVLPNKAPTYDTVLQMEY 355
Qy 388 LEAFVKESLRDPPSPYANRECLKDEDFIPLAEPVIGRDSVINEVRIKGTVMVLPFN 447
Db 356 LDMVNETLRLFPAMLERVCKXDE-----INGMFIPKGVVVMIPSYA 400
Qy 448 INRSKFIYGEDAEERFRPERWLEDVTDLSNS-IEAPYGHQASFISGPRACFGMRFAVAEMK 506
Db 401 LHRDP-KWTEPEXELPERFSKKNKNDIDPIYTPFG-----SGPRNCIGMRFALNMK 453
Qy 507 AFLFVTLRRVQFEP 520
Db 454 LALIRVLQNFSPKP 467

RESULT 10
ABB08079
ID ABB08079 standard; protein; 527 AA.
XX
AC ABB08079;
XX
DT 10-SEP-2002 (first entry)
XX
DE Maize cytochrome P450, CYP72A1.
XX
KW Maize; cytochrome P450; CYP72A1; CYP92A1; plant; gene therapy; antidote.
XX
OS Zea mays.
XX
XX Location/Qualifiers
XX Key 200..527
XX Peptide /note= "this fragment is not provided in the protein
XX sequence indicated in the sequence listing"
XX
XX US6380465-B1.
XX
XX 30-APR-2002.
XX
XX 11-JUL-1999; 99US-00351229.
XX
XX 12-JUL-1998; 98US-0092596P.
XX (KENT) UNIV KENTUCKY RES FOUND.
XX Barrett M;
XX
XX WPI; 2002-470227/50.
XX N-PSDB; ABL60758.
XX
XX Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1,
XX useful for the production of plants resistant to heterocyclic,
XX sulfonylurea and substituted urea herbicides and organophosphate
XX insecticides e.g. Classic and Pursuit.
XX
XX Claim 13; Fig 1; 21pp; English.
XX
XX The invention relates to isolated nucleic acid molecules encoding Zea
XX mays (maize) cytochrome P450 designated CYP72A1 and CYP92A1. The
XX polypeptides can be expressed by standard recombinant methodology. The
XX nucleic acids may be used in genetic engineering protocols to transform
XX plants and other eukaryotes e.g. yeast, maize (especially), soybean,
XX beet, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,
XX rose, carnation, gerbera, carrot, tomato, lettuce, chicory, pepper, melon
XX and cabbage. They may be used in this way to confer protection against
XX heterocyclic herbicides, sulfonylurea herbicides, substituted urea
XX herbicides and/or organophosphate insecticides. The present sequence
XX represents the Zea mays CYP72A1
XX
SQ Sequence 527 AA;

Query Match 13.8%; Score 392.5; DB 5; Length 527;
Best Local Similarity 24.6%; Pred. No. 3.1e-25;

DT 07-OCT-2002 (first entry)
 XX Cytochrome P450 CYP3A7 and CYP3A1 fusion protein.
 DE
 XX
 KW CYP3A7; CYP3A1; CYP3A2; CYP3A4; CYP3A5; CYP3A7;
 KW drug metabolism; drug design; drug screening.
 XX
 OS Homo sapiens.
 XX
 PN WO20024213-A1.
 XX
 PD 06-JUN-2002.
 XX
 XX 28-NOV-2001; 2001WO-SB002631.
 XX
 XX 28-NOV-2000; 2000SE-00004366.
 PR 11-JUN-2001; 2001SE-00002061.
 XX
 XX (ZAPH/) ZAPHIROPOULOS P G.
 PA (FINT/) FINTA C.
 XX
 XX Zaphiropoulos PG, Finta C;
 PI
 XX WPI; 2002-557532/59.
 DR N-ESDB; ABK97643.
 XX
 PT Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
 PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
 PT metabolism, in drug design and drug screening.
 XX
 PS Claim 1; Page 73-75; 131pp; English.
 XX
 CC The invention describes a cytochrome P450 protein (I) in which CYP3A43
 CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
 CC fragments, variants and multiples of (I) having essentially the same
 CC characteristics (I) is useful as a medicament, and for evaluating drug
 CC metabolism, in drug design, and drug screening, and in tests for
 CC adjusting the dose of drugs. This is the amino acid sequence of a novel
 CC cytochrome P450 protein
 XX
 SQ Sequence 535 AA;

 Query Match 13.5%; Score 385.5; DB 5; Length 535;
 Best Local Similarity 29.0%; Pred. No. 1.3e-24;
 Matches 157; Conservative 79; Mismatches 151; Indels 115; Gaps 26;

 QY 13 LAAPSW--ASIAFSLYL-APRRSSLYLQ-PNHTNY-FTGNFLDLSARTG-----E 61
 DB 7 LAVETWLLAVSLILLYLGRTRHGLPKLGIPOPTPLPLGN--ALGFRKGYWTFDME 63

 QY 62 BHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHVM-KEAYD-----YKPGMAARVL 113
 DB 64 CYKRYKRWG-----IYDCQPMALAITDPDMIKTVLVKCYSVFTNRRPFGVPVGMKNAI 118

 QY 114 RIATGDGVVTAAGBAHRHRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDWA 173
 DB 119 SI-----AEDEWKRISLLSPTTSCKLKEMVPIIAQYGDVLVNLREA----- 164

 QY 174 VGESAGEKKATRLT-EGVDVQWGVGRATLDVWALAGFDYKSLQNKTNLYVAVGLT 232
 DB 165 -----ETGRFVTLKHVFGAYSMDVITS*SGVSDSLNNPQD-----PVENT 207

 QY 233 D---GFAPTLDLDFKAIMWDFVPYERTWKRRIEPLTQGLAVS---RRVGIELMEQKKQAV 286
 DB 208 KKLRFNP-LDPFVLSIKVP-PFL-----TPILEALNITVPRKVISFLTCKSVKQIK 257

 QY 287 LGSASDAQVDKDVQGR--DILSLVLRANTAAANLPESQKLSDEVLAAQSNLLFAGYETS 344
 DB 258 EGRLE-----KETQKRVDFLQMLDSQNSKDSETHKALSDELMAQSIIFIFAGYETT 310

 QY 345 STVLTWTHRLSDEKAVQDKLEBICQIDTDM-----PTLDELNALPYLEAFYKESLRLD 399
 DB 311 SSVLSFIITIELATHPDVQCKVQKE---IDTVLPNKAPPTVDTVLQLELYDMVNVNLRLP 367

400 PPSVYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMMLPLFNINRSKFIYGEDA 459
 DB 368 PYAMRLERVCKDVE-----INGMFIKGVVVVWIPSYVLHDDP-KYWTEP 411

 QY 460 EBFRRPERMLEDDVTDLSNS-IBAPYHQASFTISGRACPGWFAVAEMKAFILVTLRRVQF 518
 DB 412 EKFLPERFSKKNKNDIPYITPFG-----SGPRNCIGMRFALVNMKALALVRVLFQNSF 465

 QY 519 EP 520
 DB 466 KP 467

 RESULT 13
 AAE26192
 ID AAE26192 standard; protein; 502 AA.
 XX
 AC AAE26192;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 XX Human cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) protein.
 DE
 XX Human; cytochrome P450; subfamily IIIA; polypeptide 5 isogene; CYP3A5;
 KW drug screening; polymorphism; haplotype; drug metabolising disorder;
 KW gene therapy.
 KM
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 PH
 FT Misc-difference 30 /note= "This amino acid changes to Tyr due to single
 FT nucleotide polymorphism"
 FT Misc-difference 100 /note= "This amino acid changes to Tyr due to single
 FT nucleotide polymorphism"
 FT
 XX WO200246209-A2.
 PN
 XX
 PD 13-JUN-2002.
 XX
 XX 07-DEC-2001; 2001WO-US047218.
 PF
 XX 08-DEC-2000; 2000US-0254367P.
 PR 03-MAY-2001; 2001US-0288470P.
 XX
 XX (GENA-) GENAISSANCE PHARM INC.
 XX
 XX Anastasio AE, Han J, Klien SE, Rounds E;
 PI
 XX WPI; 2002-636448/69.
 XX
 DR N-ESDB; AAD43350, AAD43351.
 XX
 XX Novel isolated polynucleotide which is a polymorphic variant of
 PT cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) gene useful for
 PT expressing CYP3A5 protein isoform used in drug screening techniques.
 XX
 PS Claim 29; Fig 3; 127pp; English.
 XX
 CC The invention relates to isolated polynucleotide having cytochrome P450,
 CC subfamily IIIA, polypeptide 5 isogene (CYP3A5). The invention is useful
 CC for screening drugs. The invention is useful for studying expression and
 CC function of CYP3A5 and expressing CYP3A5 protein for use in screening for
 CC candidate drugs to treat diseases related to CYP3A5 activity. The
 CC polymorphism and haplotype data is useful for validating whether CYP3A5
 CC is a suitable target for drugs to treat drug metabolising disorders.
 CC screening for such drugs and reducing bias in clinical trials of such
 CC drugs. The invention is also useful for therapeutic purposes. The
 CC invention is useful in studying the effect of variation on the biological
 CC activity of CYP3A5 as well as on the binding affinity of candidate drugs
 CC to CYP3A5, or for studying the enzymatic properties of such CYP3A5
 CC variants using these candidate drugs as substrate. The invention is

CC useful in gene therapy. The present sequence is human CYP3A5 protein

XX Sequence 502 AA;

Query Match 13.5%; Score 385; DB 5; Length 502;
 Best Local Similarity 28.1%; Pred. No. 1.3e-24;
 Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYLQ-PNHTNY-FTGNFLDILSARTG-----E 61
 Db 7 LAVETWLLAVSLVLLYLYGTRTHGLFKRLGIPGTPPLLGN---VLSYRQGLWKEDTE 63

QY 62 EHAKYREKYGSTRFAGIAGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATG-- 118
 Db 64 CYKYGKMGWGT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111

QY 119 ----DGVVTAEGEAHRRHRIWIPSLSAQAVKSVPIFLEKGMELVDKMDAEAKDMAV 174
 Db 112 GFMKSAISLAEDDEWKIRSLSPFTSGKLEMPFIIAQYGDVLRVLRRE-AEKG--- 167

QY 175 GESAGEKKATRLTEGVVDKVGVRATLDVVALAGFDYKSDSLQN-----KTNELYVAF 228
 Db 168 -----KPVTLKDIFGAYSDMVTGTSGFVNIIDSLNNPQDPFVSTKFLKF 213

QY 229 VGLTDFGAPTLDSFKAIMDVPYFRTMKRRHEIPLTQGLAVS--RRVGLMEQKKQAV 286
 Db 214 -----GFLDPL-FLSIILFPFL-----TFVFEALNVSLFPKDTINFLSKS--- 252

QY 287 LGSASDAQVDKQVQGRDILSLVRANIANLPESQKLSDEEVLAAQISNLLFAGYETSST 346
 Db 253 VNRMKSRNLNDKQKRLDFQLMIDSONSKETESHKALSDELAQAQSIIFIFAGYETSS 312

QY 347 VLTWMPHRLSEDKAVQDKLREIQT--DTDMPTLDELNALPYLEAFVKESLRDPPSPY 404
 Db 313 VLSFTLYELATHPDVQOKLQKEIDAVLPNKAPPTYDAVVQMEYLDVMVNETLRLFPVAIR 372

QY 405 ANRECKDDEDFIPLAEFVIGRGSVINEVRIITKGMWMLPLFNINRSFIYGEDAEERFP 464
 Db 373 LEPTCKRDVE-----INGVFIKGMVNVPIYALHNDP-KWTPEPERFP 416

QY 465 ERWLEDVTDLSNS-IEAPYGHQASFTISGPRACFGNRFVAEMKAPLFTVLRVQFEP 520
 Db 417 ERSKSK-KDSIDPIYITPFG-----TGPRNCIGMRFALNMKALIRVLQNFSPK 466

RESULT 14
 ADE57186
 AC ADE57186 standard; protein; 502 AA.
 AC ADE57186;
 DT 29-JAN-2004 (first entry)
 DE Human Protein P20815, SEQ ID NO 3047.
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS
 XX WO2003016475-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 FF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PR (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WPI; 2003-268312/26.
 XX GENBANK; F20815.
 DR
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PS Claim 1; Page; 101pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 502 AA;

Query Match 13.5%; Score 385; DB 7; Length 502;
 Best Local Similarity 28.1%; Pred. No. 1.3e-24;
 Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYLQ-PNHTNY-FTGNFLDILSARTG-----E 61
 Db 7 LAVETWLLAVSLVLLYLYGTRTHGLFKRLGIPGTPPLLGN---VLSYRQGLWKEDTE 63

QY 62 EHAKYREKYGSTRFAGIAGAPVLNSTDPKVFNVHM-KEAYDYPKPGMAARVLRATG-- 118
 Db 64 CYKYGKMGWGT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111

QY 119 ----DGVVTAEGEAHRRHRIWIPSLSAQAVKSVPIFLEKGMELVDKMDAEAKDMAV 174
 Db 112 GFMKSAISLAEDDEWKIRSLSPFTSGKLEMPFIIAQYGDVLRVLRRE-AEKG--- 167

QY 175 GESAGEKKATRLTEGVVDKVGVRATLDVVALAGFDYKSDSLQN-----KTNELYVAF 228
 Db 168 -----KPVTLKDIFGAYSDMVTGTSGFVNIIDSLNNPQDPFVSTKFLKF 213

QY 229 VGLTDFGAPTLDSFKAIMDVPYFRTMKRRHEIPLTQGLAVS--RRVGLMEQKKQAV 286
 Db 214 -----GFLDPL-FLSIILFPFL-----TFVFEALNVSLFPKDTINFLSKS--- 252

QY 287 LGSASDAQVDKQVQGRDILSLVRANIANLPESQKLSDEEVLAAQISNLLFAGYETSST 346
 Db 253 VNRMKSRNLNDKQKRLDFQLMIDSONSKETESHKALSDELAQAQSIIFIFAGYETSS 312

QY 347 VLTWMPHRLSEDKAVQDKLREIQT--DTDMPTLDELNALPYLEAFVKESLRDPPSPY 404
 Db 313 VLSFTLYELATHPDVQOKLQKEIDAVLPNKAPPTYDAVVQMEYLDVMVNETLRLFPVAIR 372

QY 405 ANRECKLDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFICYDEAEERFP 464
Db 373 LERTCKRQVE-----INGVFIPKGSWVIPTALHHP-KYWTPEERFP 416
QY 465 ERWLEDVTDLSNS-IBAPYGHQASFTSGPRACFGWRFAVAEMKAFLEFVTLRRVQFEP 520
Db 417 ERSKX-KDSIDPVIYTPFG-----TGPRNCIGMRFAFMNMKALIRVLQNFSPKP 466

RESULT 15
ADE57190
ID ADE57190 standard; protein; 502 AA.
XX AC ADE57190;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein P20815, SEQ ID NO 3051.
XX KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX WO2003016475-A2.
XX PN 27-FEB-2003.
XX PD 14-AUG-2002; 2002WO-US025765.
XX PF 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX DR GENBANK; P20815.
XX DR

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
XX The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 502 AA;
QY Query Match 13.5%; Score 385; DB 7; Length 502;
Db Best Local Similarity 28.1%; Pred. No. 1.3e-24;
QY Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;
Db 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61
Db 7 LAVETWLLAVSLVLLVLYGTRTHGLFRGLGIPGPTPLPLGN---VLSYQGLWKPKDTE 63
QY 62 ERAKYREKYGSTIRFAGTAGAPVLNSTDPKVFNVHM--KEAYDYPKPGMAARVLRIATG-- 118
Db 64 CYKKGKMGWT--YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRRSLGPV 111
QY 119 -----DGVTTAEAEAKHRRIMIPSLSAQAVKSMVPIFEKGMELVDKMDAAEKDQAV 174
Db 112 GFMKSAISLAEDDEEWKRISLLSPFTSGKLKEMPPIIAQYGDVLVRLRRE-AEKG--- 167
QY 175 GESAGEKKATRLTEGVDVKDWVGRATLDVVALAGFDYKSDSLQN-----KTNELYVAF 228
Db 168 -----KPVTLKIDIFGAYSMVDVITGTSFGVNIDSLNNPQDPFVESTKKPKF 213
QY 229 VGTGDPAPTLDSFKAIMWDFVYPTWKRRHEIPLTQGLAVS---RRVGIELMEQKKQAV 286
Db 214 -----GFLDPL-FLSILFPFL-----TFVFEALNVSLFPKDTINFLSKS---- 252
QY 287 LGSASDOAVDKDQVQGRDILSLVRANTAAANPEQSKLSDEVLQAISNLLFAGYETSST 346
Db 253 VNRMKSRINDKQKHLDFLQMLDSQNSKETESHKALSDLELAQSIIFIFAGYETTS 312
QY 347 VLTWFMHRLSEDKAVQDKLRBEICQI--DTDMPTLDELNALPYLEAFVKESIRLDPPSPY 404
Db 313 VLSFTLYELATHPDVQOKLQKEIDAVLPNKAPPTYDAVVQMEYLDVMVNETIRLFPVAIR 372
QY 405 ANRECKLDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFICYDEAEERFP 464
Db 373 LERTCKRQVE-----INGVFIPKGSWVIPTALHHP-KYWTPEERFP 416
QY 465 ERWLEDVTDLSNS-IBAPYGHQASFTSGPRACFGWRFAVAEMKAFLEFVTLRRVQFEP 520
Db 417 ERSKX-KDSIDPVIYTPFG-----TGPRNCIGMRFAFMNMKALIRVLQNFSPKP 466
Search completed: April 2, 2004, 13:58:35
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 10:03:02 ; Search time 717 Seconds

(without alignments)
10031.675 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932
Sequence: 1 gaattcgccagcagccagccaccc.....agccgctcgtgcagattc 1932

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1932	100.0	1932	14	US-10-066-007-2
2	353.2	18.3	3969	14	US-10-066-007-4
3	49.8	2.6	467	12	US-10-424-599-20949
4	49.4	2.6	1501	15	US-10-264-237-392
5	49.2	2.5	368	12	US-10-424-599-1793
6	48.4	2.5	657	12	US-10-424-599-25496
7	48.4	2.5	1512	15	US-10-313-863A-55
8	48.4	2.5	2011	9	US-09-880-107-1586
9	48.4	2.5	2759	14	US-10-146-575-1
10	48.4	2.5	2768	14	US-10-268-822-14
11	48.2	2.5	2849	9	US-09-880-107-2110
12	48.2	2.5	312	9	US-09-960-357-4611
13	47.6	2.5	3059	9	US-09-925-297-268
14	47.4	2.5	628	12	US-10-424-599-71313
15	47.2	2.4	1228	12	US-10-424-599-32438

C 16	47	2.4	218	9	US-09-933-797-162	Sequence 162, App
C 17	47	2.4	685	12	US-10-424-599-8656	Sequence 8656, App
C 18	46.8	2.4	709	12	US-10-424-599-75080	Sequence 75080, A
C 19	46.8	2.4	1151	12	US-10-424-599-23175	Sequence 23175, A
C 20	46.8	2.4	1712	9	US-09-981-876-106	Sequence 106, App
C 21	46.8	2.4	1712	10	US-09-148-545-106	Sequence 106, App
C 22	46.8	2.4	1822	9	US-09-981-876-105	Sequence 105, App
C 23	46.8	2.4	1822	10	US-09-148-545-105	Sequence 105, App
C 24	46.6	2.4	1091	12	US-10-424-599-33549	Sequence 33549, A
C 25	46.4	2.4	1419	12	US-10-424-599-72966	Sequence 72966, A
C 26	46.2	2.4	672	12	US-10-424-599-99744	Sequence 106636, A
C 27	46.2	2.4	710	12	US-10-424-599-99744	Sequence 99744, A
C 28	46.2	2.4	1194	12	US-10-424-599-15121	Sequence 15121, A
C 29	46	2.4	266	12	US-10-424-599-29930	Sequence 29930, A
C 30	46	2.4	414	12	US-10-424-599-44827	Sequence 44827, A
C 31	46	2.4	940	12	US-10-424-599-94220	Sequence 94220, A
C 32	45.8	2.4	592	14	US-10-066-543-950	Sequence 111262, A
C 33	45.8	2.4	592	14	US-10-066-543-950	Sequence 950, App
C 34	45.8	2.4	1498	12	US-10-424-599-54757	Sequence 54757, App
C 35	45.8	2.4	1498	12	US-10-424-599-54758	Sequence 54758, A
C 36	45.6	2.4	2301	12	US-10-424-599-138418	Sequence 138418, A
C 37	45.6	2.4	447	12	US-10-424-599-142657	Sequence 142657, A
C 38	45.6	2.4	583	12	US-10-424-599-52047	Sequence 52047, A
C 39	45.6	2.4	705	12	US-10-424-599-137273	Sequence 137273, A
C 40	45.6	2.4	733	12	US-10-424-599-118266	Sequence 118266, A
C 41	45.4	2.3	558	12	US-10-424-599-67477	Sequence 67477, A
C 42	45.4	2.3	662	12	US-10-424-599-80087	Sequence 80087, A
C 43	45.4	2.3	767	12	US-10-424-599-43674	Sequence 43674, A
C 44	45.4	2.3	1353	15	US-10-369-493-141849	Sequence 141849, A
C 45	45.4	2.3	2508	12	US-10-424-599-141849	Sequence 141849, A

ALIGNMENTS

RESULT 1
US-10-066-007-2
Sequence 2, Application US/10066007

PUBLICATION NO. US2003077691A1

GENERAL INFORMATION:

Query Match 100.0%; Score 1932; DB 14; Length 1932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 GAATTCGCGCAGGAGCCACCTTCTTCATATGTCATCTTGTGCTGCACAGTG 60

PRIOR APPLICATION NUMBER: EP 99104668.1
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: EP 00101666.6
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 4

LENGTH: 3969

TYPE: DNA

ORGANISM: Phaffia rhodozyma

FEATURE:

NAME/KEY: 5'UTR

LOCATION: (517)..(518)

NAME/KEY: intron

LOCATION: (784)..(898)

NAME/KEY: intron

LOCATION: (1016)..(1087)

NAME/KEY: intron

LOCATION: (1180)..(1302)

NAME/KEY: intron

LOCATION: (1518)..(1600)

NAME/KEY: intron

LOCATION: (1635)..(1723)

NAME/KEY: intron

LOCATION: (1867)..(1939)

NAME/KEY: intron

LOCATION: (2000)..(2081)

NAME/KEY: intron

LOCATION: (2182)..(2257)

NAME/KEY: intron

LOCATION: (2355)..(2431)

NAME/KEY: intron

LOCATION: (2543)..(2618)

NAME/KEY: intron

LOCATION: (2653)..(2742)

NAME/KEY: intron

LOCATION: (2815)..(2962)

NAME/KEY: intron

LOCATION: (3051)..(3113)

NAME/KEY: intron

LOCATION: (3172)..(3247)

NAME/KEY: intron

LOCATION: (3322)..(3398)

NAME/KEY: intron

LOCATION: (3424)..(3513)

NAME/KEY: polyA site

LOCATION: (3865)..(3866)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

NAME/KEY: intron

LOCATION: (653)..(734)

Db 3748 ACTAGGACAAATGCTTCTTTGTATGATGATTGTTCTCATACCCGGCAGGCGCTATGACT 3807
Qy 1814 TTACGCTCTATGCTGCTCTGACTCTCTTCTTACCTATATATTTCCATCCG 1871
Db 3808 TCTACGCTGCTATGCTGCTCTGACTCTCTTCTTACCTATATATTTCCATCCG 3865

RESULT 3

US-10-424-599-20949

Sequence 20949, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 39-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 20949

LENGTH: 467

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_118922C.1

US-10-424-599-20949

Query Match 2.6%; Score 49.8; DB 12; Length 467;

Best Local Similarity 96.2%; Pred. No. 0.00018; Indels 0; Gaps 0;

Matches 51; Conservative 0; Mismatches 2;

Qy 1861 TATTCCATCCGAAAAA..... 1913

Db 401 TATTCAATCCAAAAA..... 453

RESULT 4

US-10-264-237-392

Sequence 392, Application US/10264237

Publication No. US20040009491A1

GENERAL INFORMATION:

APPLICANT: Birse et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P131P1

CURRENT APPLICATION NUMBER: US/10/264,237

CURRENT FILING DATE: 2002-10-04

PRIOR APPLICATION NUMBER: PCT/US01/16450

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: US 60/205,515

PRIOR FILING DATE: 2000-05-19

NUMBER OF SEQ ID NOS: 2876

SOFTWARE: Patent in Ver. 3.1

SEQ ID NO 392

LENGTH: 1501

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (7)..(7)

OTHER INFORMATION: n equals a.t.g, or c

US-10-264-237-392

Query Match 2.6%; Score 49.4; DB 15; Length 1501;

Best Local Similarity 63.2%; Pred. No. 0.0005;

Matches 74; Conservative 1; Mismatches 42; Indels 0; Gaps 0;

Qy 1785 TTTCCTCATACCCGGCAGGCGCTATGACTTCTACGTCGTCTCGTCTCTCGTCTCTC 1844

Db 1370 TTTCCTTGCTTGCCTCTTTGTGCACAGACTGCATAAGTTGTGCAGCTTGCATTC 1429


```

; SEQ ID NO 1586
; LENGTH: 2011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00003
US-09-880-107-1586

Query Match      2.5%; Score 48.4; DB 9; Length 2011;
Best Local Similarity 50.7%; Pred. No. 0.0012;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGAGGTAAGTCTGATCAGATCAGTAACTGTTATTTGCTGGATATG 1056
DB 926 AAGCTCTGTCGATGAGAGGTAAGTCTGATCAGATCAGTAACTGTTATTTGCTGGATATG 985
QY 1057 AAACTTCTTCGACAGTCTTGACATGATGTTTCCCGACTCTCAGAGACAAAGCCGTTTC 1116
DB 986 AAACCCAGCAGGAGTCTCTCTCTTCATTATGATGATGATGATGATGATGATGATGATG 1045
QY 1117 AGGATAAACTTCGAGAGAAATTTGTCAGATCAGACGATGATGATGATGATGATGATG 1176
DB 1046 AGCAGAACTGCAGGAGGAAATTTGATGATGATGATGATGATGATGATGATGATGATG 1105
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
DB 1106 ATCTGTGCTACAGATGGAGTAATCTTGACATGGTGGTGAATGAACGCTCAGATTATCC 1165
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGGATGAGACTTCA 1278
DB 1166 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA 1213

RESULT 9
US-10-146-575-1
; Sequence 1, Application US/10146575
; Publication No. US20030059800A1
; GENERAL INFORMATION:
; APPLICANT: Lichter, Jay
; APPLICANT: Guido, Marco
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-12P
; CURRENT APPLICATION NUMBER: US/10/146,575
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US/09/144,367
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2759
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)...(1581)
; OTHER INFORMATION: Human CYP3A4 cDNA reference sequence
US-10-146-575-1

Query Match      2.5%; Score 48.4; DB 14; Length 2759;
Best Local Similarity 50.7%; Pred. No. 0.0015;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGAGGTAAGTCTGATCAGATCAGTAACTGTTATTTGCTGGATATG 1056
DB 932 AAGCTCTGTCGATGAGAGGTAAGTCTGATCAGATCAGTAACTGTTATTTGCTGGATATG 991
QY 1057 AAACTTCTTCGACAGTCTTGACATGATGTTTCCCGACTCTCAGAGACAAAGCCGTTTC 1116
DB 992 AAACCCAGCAGGAGTCTCTCTCTTCATTATGATGATGATGATGATGATGATGATGATG 1051
QY 1117 AGGATAAACTTCGAGAGAAATTTGTCAGATCAGACGATGATGATGATGATGATGATGATG 1176
DB 1052 ACAGAAACTGCAGGAGGAAATTTGATGATGATGATGATGATGATGATGATGATGATG 1111

```

```

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
DB 1112 ATACTGTCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATCC 1171
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGGATGAGACTTCA 1278
DB 1172 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA 1219

RESULT 10
US-10-268-822-14
; Sequence 14, Application US/10268822
; Publication No. US20030150004A1
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Wei, Ping
; APPLICANT: Chua, Steven
; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob
; TITLE OF INVENTION: Metabolism
; FILE REFERENCE: P02729US2
; CURRENT APPLICATION NUMBER: US/10/268,822
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US 01/29672
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 10/219,590
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 14
; LENGTH: 2768
; TYPE: DNA
; ORGANISM: Human
US-10-268-822-14

Query Match      2.5%; Score 48.4; DB 14; Length 2768;
Best Local Similarity 50.7%; Pred. No. 0.0015;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGATGAGAGGTAAGTCTGATCAGATCAGTAACTGTTATTTGCTGGATATG 1056
DB 967 AAGCTCTGTCGATGAGAGGTAAGTCTGATCAGATCAGTAACTGTTATTTGCTGGATATG 1026
QY 1057 AAACTTCTTCGACAGTCTTGACATGATGTTTCCCGACTCTCAGAGACAAAGCCGTTTC 1116
DB 1027 AAACCCAGCAGGAGTCTCTCTCTTCATTATGATGATGATGATGATGATGATGATGATG 1086
QY 1117 AGGATAAACTTCGAGAGAAATTTGTGATGATGATGATGATGATGATGATGATGATGATG 1176
DB 1087 AGCAGAACTGCAGAGGAAATTTGATGATGATGATGATGATGATGATGATGATGATG 1146
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
DB 1147 ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATCC 1206
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGGATGAGACTTCA 1278
DB 1207 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA 1254

RESULT 11
US-09-880-107-2110
; Sequence 2110, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14

```

	Query Match	2.5%	Score 48.2	DB 9	Length 312
	Beat Local Similarity	55.8%	Prod. No. 0.00044		
	Matches 92	Conservative	0	Mismatches 73	Indels 0
				Gaps 0	
QY	1745	GAATGTGTGACTAGGACAATGCCCTTCTTTGTTATCGATTTGTTTCTATACCCGGGAGCC	1804		
DB	175	GAATCCAGGAGGAAACACTGAGACCTTAGAGATTAAATGTAGTCTAAACACCTAGCCATA	116		
QY	1805	GCATGACTTCTACGCTCTCTATCGCTCTGCACTCTCTCTTCAACCTATATATTATT	1864		
DB	115	TGATAGACTGATATCATGTATCACTGGACTTTTCTGCTTTAAATAAAACCAATACGANT	56		

	Query Match	2.5%	Score 47.4;	DB 12;	Length 628;
	Best Local Similarity	89.5%;	Pred. No. 0.0012;		
	Matches 51; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1858	TATTATTCCTCCGAGAAAAA	AAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAGCGCCTCTGAG	1914
nB	61	TATTACGTCCAGCAAAAAA	AAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAGCGCCTCTAG	5

RESULT 15

```
US-10-424-599-32438/c
; Sequence 32438, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 32438
; LENGTH: 1228
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129297C.1
US-10-424-599-32438

Query Match          2.4%; Score 47.2; DB 12; Length 1228;
Best Local Similarity 86.7%; Pred. No. 0.0021; 8; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 8;

Oy 1855 ATATATTATTCCATCCGAAAAAAGCGCCGCTCGAG 1914
    |||||
Db 68 ATGTGTCTTCCCGCCCAAAAAAAGCGCCGCTCTAG 9

Search completed: March 31, 2004, 12:47:06
Job time : 723 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 54 Seconds
(without alignments)
2914.425 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852
Sequence: 1 MFLVLTGALGLAAPSWS.....RIVGREKEGYQNRLOQKPE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	3	AB08713 Amino aci
2	398.5	14.0	503	6	ABU57260 Human chy
3	398.5	14.0	503	6	ABR82024 Human cyt
4	396.5	13.9	502	7	ADD48378 Human Pro
5	396.5	13.9	503	2	AAR72363 Human cyt
6	396.5	13.9	503	2	AAR81464 Human der
7	396.5	13.9	503	2	AAR93170 Human cyt
8	396.5	13.9	503	2	AA05202 Human CYP
9	394	13.8	503	5	ABG68753 Cytochrom
10	392.5	13.8	527	5	ABO8079 Maize cyt
11	385.5	13.5	503	5	AAE22852 Human cyt
12	385.5	13.5	535	5	ABG68747 Cytochrom
13	395	13.5	502	5	AAE26192 Human cyt
14	385	13.5	502	7	ABE57186 Human Pro
15	385	13.5	502	7	ABE57190 Human Pro
16	385	13.5	502	7	ABE663919 Human Pro
17	382	13.4	454	5	ABG68754 Cytochrom
18	373.5	13.1	454	6	ADA48312 Rice prot
19	370	13.0	504	7	ABE57188 Rat Prote
20	370	13.0	504	7	ABE63917 Rat Prote
21	370	13.0	504	7	ABE57184 Rat Prote
22	367	12.9	520	3	ABE23917 Arabidops
23	365	12.8	520	6	ABU54571 Human NOV
24	363.5	12.7	512	5	AAU97096 Absciscic
25	362.5	12.7	465	3	AA054205 Arabidops

26	362.5	12.7	475	3	AA054204	Arabidops
27	362.5	12.7	523	3	AA054203	Arabidops
28	360	12.6	512	2	AA055998	Vicia sat
29	356.5	12.5	475	3	AA024685	Arabidops
30	356.5	12.5	522	3	AA024684	Arabidops
31	356.5	12.5	527	3	AA024683	Arabidops
32	354	12.4	562	4	AAU30319	Novel hum
33	351	12.3	512	5	AAU97097	Abciscic
34	351	12.3	1115	6	AA016061	Gibberell
35	350.5	12.3	1054	6	AA016055	Bacillus
36	350.5	12.3	1054	6	AA016056	Bacillus
37	349	12.2	508	5	ABR07528	Human dru
38	348.5	12.2	511	7	AD60554	Human Pro
39	348.5	12.2	546	4	AAU14298	Human nov
40	347.5	12.2	524	4	AAU14534	Human nov
41	347.5	12.2	524	4	AAU14536	Human nov
42	347.5	12.2	524	4	AAU14535	Human nov
43	346.5	12.1	511	7	AD60552	Rat Prote
44	346.5	12.1	524	3	AA056706	Membrane-
45	346.5	12.1	524	4	AA073677	Human oxi

ALIGNMENTS

RESULT 1

AA080713
ID AAB08713 standard; protein; 557 AA.

AC AAB08713;

XX AAB08713;

DT 12-SEP-2003 (revised)

DT 02-JAN-2001 (first entry)

XX Amino acid sequence of an astaxanthin synthetase polypeptide.

DE Astaxanthin synthetase; astaxanthin; beta-carotene; carotenogenic yeast;

XX antioxi-dant; cancer; colouring reagent; farmed fish; salmon.

KW Xanthophyllomyces dendrorhous.

OS EP1035206-A1.

PN 13-SEP-2000.

XX 03-MAR-2000; 2000EP-00104430.

PF 09-MAR-1999; 99EP-00104668.

PR 01-FEB-2000; 2000EP-00101666.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Hoshino T, Ojima K, Setoguchi Y;

XX WPI; 2000-559874/52.

DR N-PSDB; AAA64472, AAA64473.

XX Novel polynucleotide encoding astaxanthin synthase useful for producing

PT recombinant cells for producing astaxanthin from beta-carotene.

XX Claim 2; Page 21-23; 46pp; English.

XX The present sequence represents an astaxanthin synthetase polypeptide of

CC Phaffia rhodozyma. The enzyme is involved in the last step of the

CC astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.

CC rhodozyma is a carotenogenic yeast strain. The astaxanthin synthetase

CC polynucleotides and polypeptides are useful for producing astaxanthin.

CC Astaxanthin is an antioxidant which may be used to protect living cells

CC against diseases such as cancer. Astaxanthin is also used as a colouring

CC reagent, e.g. in farmed fish like salmon to impart an orange-red

XX coloration. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 557 AA;

RESULT 3
 ID ABR82024 standard; protein; 503 AA.
 XX ABR82024;
 AC ABR82024;
 DT 22-SEP-2003 (first entry)
 DE Human cytochrome P450 3A4 amino acid sequence.
 XX Human; protein array; protein moiety; phenotype; drug discovery;
 KW naturally occurring variant; pharmacogenomic; diagnostic;
 KW parallel analysis; tumour suppressor; p53; cytochrome P450.
 XX Homo sapiens.
 XX WO2003048768-A2.
 XX 12-JUN-2003.
 XX 05-DEC-2002; 2002WO-GB005499.
 XX 05-DEC-2001; 2001US-0335806P.
 XX 16-SEP-2002; 2002US-0410815P.
 XX (SENS-) SENSE PROTEOMIC LTD.
 PA Boutell JW, Godber BLJ, Hart DJ, Blackburn JD;
 FI WPI; 2003-569063/53.
 DR N-PSDB; ACF06054.
 XX New protein array, useful for determining the phenotype of a naturally
 PT occurring variant of a DNA sequence of interest, comprises a surface upon
 PT which at least two protein moieties are deposited.
 XX Example 5; Fig 11B; 84pp; English.
 XX The present invention describes a protein array comprising a surface upon
 CC which at least two protein moieties are deposited at spatially defined
 CC locations, where the protein moieties are naturally occurring variants of
 CC a DNA sequence of interest. Also described: (1) making a protein array;
 CC (2) screening a set of protein moieties for molecules that interact with
 CC one or more proteins; and (3) simultaneously determining the relative
 CC properties of members of a set of protein moieties. The protein array can
 CC be used for determining the phenotype of a naturally occurring variant of
 CC a DNA sequence of interest. The protein array is useful for drug
 CC discovery, pharmacogenomics and diagnostics. The protein array allows the
 CC parallel analysis of closely related proteins with a sensitivity that is
 CC at least comparable to existing methods, if not better, with small
 CC volumes of potentially expensive ligands, and in a quantitative,
 CC comparative, functional analysis manner not previously possible. ACF06000
 CC to ACF06056 and ABR81975 to ABR82025 represent sequences used in the
 CC exemplification of the present invention
 XX Sequence 503 AA;
 SQ
 Query Match 14.0%; Score 398.5; DB 6; Length 503;
 Best Local Similarity 28.7%; Pred. No. 8.6e-26;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 13 LAATSW--ASTAFSLYLAPRS--SLYNLOG--PNHTNY--FTGNFLILSARGT-----E 61
 7 LAMETLLAVSLVLLLYLTHSHGLFKLGIQPTPLPFLGN---ILSYHKGFCNFDME 63
 62 BHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNHVM--KEAYDY---PKPQMAARVLRIAT 117
 64 CHKTKGWGF---YDG--QQVLAIIDPDIMKTVLVKCYSVFTNRPPGPGVFMKSA- 117
 118 GDGVVTAEGEAKHRRIRIMISLSAQAVKGNVPIFEKGMELVDKMMEDAAEKDMVAGES 177
 118 ---ISIADEEWKRLRLSLSPFTSGKLKEMVPIAQYGDVLRNLRREA----- 164

178 AGEKATRLT-EGVDVKDWVGRATLDVMAAGFDYKSDSLQKNTNELYVAFLVGLTGDA 236
 165 -----ETGKPTLKDFGAYSMVITSTSGVNIIDSLNPD----- 201
 237 PILDSEKAIM-WDFV-PYERTMK-RHEIPLTQGLAV---SRRVGIELMEQKQAVLGSA 290
 202 PFVENTKLLRDFDLDPFLSITVFPFLIPLELVNICVFPREVTNFKSVKR-----M 256
 291 SDQAVDKKDVQGRDILSLVRANITANLPESQKLSDEVLQAISNLLFAGYETSTVLTW 350
 257 KESRLEDTQKHRVDFLQMLMIDSONSKETESHKALSLELVAQSIIFIFAGYETSTVLSF 316
 351 MHRILSEKAVQDKLREICQI--DTDMPTLDLNAFLYEAFAVKESIRLDPSPYANRE 408
 317 IMVELATHPDVQOQKLEIDAVLPKAPPTDVLQMEYLDVWVNEILRLFPIMRLERV 376
 409 CLKDEDFIPLABEVIGRDGSVINEVRIITKGTVMPLFLFNINRSKFIYGEDAEERPRPWL 468
 377 CKKQVE-----INGMFIKGVVVMIPSYALHRDP-KYWTEPEKTLPERFS 420
 469 EDVTDLSNS-IEAPYGHQASFISSGRACFGWRFAVAEMKAPLFVTLRRVQFEP 520
 421 KKKKNIDPIYIYTPFG-----SGPRNCIGMRFAALMNMKLLALIRVLQNFSEKP 467

RESULT 4
 ADD48378
 ID ADD48378 standard; protein; 502 AA.
 XX ADD48378;
 AC ADD48378;
 DT 29-JAN-2004 (first entry)
 DE Human Protein A29815, SEQ ID NO 14079.
 XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 XX WO2003016475-A2.
 XX 27-FEB-2003.
 XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 XX 01-NOV-2001; 2001US-0346382P.
 XX 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 XX (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 XX GENBANK; A29815.
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 101pp; English.
 PS The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 502 AA;

Query Match 13.9%; Score 396.5; DB 7; Length 502;
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 QY 13 LAAFSW--ASIAFFSLLYAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
 DB 6 LAMETWLLAVSLVLLYLGTHSHGLFKKLGIPGPTPLPFLGN---ILSYHKGFCFMDME 62
 QY 62 EHAKEYKYSTLRPAGIAGAPVLNSTDPKVNHYM--KEAYDY---PKPGMAARVLRIAT 117
 DB 63 CHKKYKQWGF---YDG--QQPVLAITDPDMKTVLVKECTSVFTNRPPFGVGFMSKA- 116
 QY 118 GDGVVTAEGEAHKKRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDMAYGES 177
 DB 117 ---ISIADEEWKRLRSLLSPTFTSGKLKEMVPIIAQYGDVLVRLREA----- 163
 QY 178 AGEKKAATLET-EGVDVKDWVGRATLDVMAAGFDYKSDSLQNTKNELYVAFVGLTDGFA 236
 DB 164 -----ETGKPVTLKDVFGAYSMVDITSTSGVNIIDSLNPNOD----- 200
 QY 237 PTLDSFKAIM-WDFV-PYFRTWK--RRHEIPLTQGLAV---SRRVGIELMEOKKQAVLGSA 290
 DB 201 PFVENTKLLRDFLDPPFLSTVFPFLPILEVLNLCVFPREVTNFRKSVKR-----M 255
 QY 291 SDQAVDKVQGRDILSLVLRANIANLPESQKLSDEVLQAISNLLFAGYETTSVLTW 350
 DB 256 KESRLEDTKHRVDFLQMLDMSQNSKXETESHKALSDELVAQSIIIFAGYETTSVLSF 315
 QY 351 MFLRLSDRAVDKLRREEICQI--DTPMTLDELNALPYLEAFVKESLRLDPPGPPYANRE 408
 DB 316 IYELATHDPVQKQLEEDDAVLNPKAPFTYDTVLCMEYLDVWVNETLRLPIARLERV 375
 QY 409 CLKDEDFIPLAEPVIGRDSVINEVITKGTVMVMLPLFNINRSKFTYGEDAEERPRERWL 468
 DB 376 CKKDVE-----INGMEIPKGVWVM-PSVALHRDP-KYWTPEKEKTLPERFS 419
 QY 469 EDVTDLSINS-IEAPYCHQASFTSGPRACFGWFAVEMKAEFLVTLRRVQFEP 520
 DB 420 KKNKNIDPIYTPFG-----SGPRNCIGMRPALNWKALIRVLQNFSPK 466

RESULT 5
 AAR72363
 ID AAR72363 standard; protein; 503 AA.
 XX
 AC AAR72363;
 XX
 XX
 DT 25-MAR-2003 (revised)
 DT 14-NOV-1995 (first entry)
 XX
 XX Human cytochrome P450 molecular species 3A4 protein.
 DE
 XX

KW Human cytochrome P450; amplification; PCR; primer; expression vector;
 KW yeast NADPH-P450 reductase, safety; fusion protein; metabolite;
 KW carcinogen; mutagen; liver metabolism.
 XX
 OS Homo sapiens.
 PN EP644267-A2.
 XX
 PD 22-MAR-1995.
 XX
 PF 20-JUL-1994; 94EP-00111298.
 XX
 XX 20-JUL-1993; 93JP-00201120.
 PR 21-JUL-1993; 93JP-00180246.
 PR 30-JUL-1993; 93JP-00208279.
 XX
 XX (SUMO) SUMITOMO CHEM CO LTD.
 PA
 PI Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
 DR WPI; 1995-116991/16.
 XX
 XX N-PSDB; AAQ87717.
 PT Evaluation of safety of a chemical cpd. - using recombinant Yeast
 PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
 XX
 XX Example; Page 31-33; 124pp; English.
 XX
 CC The amino acid sequence of the human cytochrome P450 species 3A4. The
 CC cDNA was amplified by PCR using the primers AAQ87743-6. The product was
 CC cloned into the yeast expression vectors pAAH8 or pAHR8 to produce the
 CC vectors p3A4 for the expression of the cytochrome P450 alone or p3A4R for
 CC co-expression with the yeast NADPH-P450 reductase. The vectors are used
 CC in a method for evaluating the safety of a chemical compound by reacting
 CC the chemical compound with recombinantly produced human cytochrome P450
 CC molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4,
 CC or their auxiliary species and variants (AAQ87718-32), and yeast NADPH-
 CC P450 reductase, either as a fused protein or in cell extracts, and
 CC analysing the resulting metabolite to assess the safety of the chemical
 CC compound. The method is useful for determining whether the chemical
 CC compound, or its metabolite, will be converted into a carcinogenic or
 CC mutagenic form through metabolism in the liver. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX
 XX Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;
 Best Local Similarity 28.7%; Pred. No. 1.3e-25;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 QY 13 LAAFSW--ASIAFFSLLYAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
 DB 7 LAMETWLLAVSLVLLYLGTHSHGLFKKLGIPGPTPLPFLGN---ILSYHKGFCFMDME 63
 QY 62 EHAKEYKYSTLRPAGIAGAPVLNSTDPKVNHYM--KEAYDY---PKPGMAARVLRIAT 117
 DB 64 CHKKYKQWGF---YDG--QQPVLAITDPDMKTVLVKECTSVFTNRPPFGVGFMSKA- 117
 QY 118 GDGVVTAEGEAHKKRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKDMAYGES 177
 DB 118 ---ISIADEEWKRLRSLLSPTFTSGKLKEMVPIIAQYGDVLVRLREA----- 164
 QY 178 AGEKKAATLET-EGVDVKDWVGRATLDVMAAGFDYKSDSLQNTKNELYVAFVGLTDGFA 236
 DB 165 -----ETGKPVTLKDVFGAYSMVDITSTSGVNIIDSLNPNOD----- 201
 QY 237 PTLDSFKAIM-WDFV-PYFRTWK--RRHEIPLTQGLAV---SRRVGIELMEOKKQAVLGSA 290
 DB 202 PFVENTKLLRDFLDPPFLSTVFPFLPILEVLNLCVFPREVTNFRKSVKR-----M 256
 QY 291 SDQAVDKVQGRDILSLVLRANIANLPESQKLSDEVLQAISNLLFAGYETTSVLTW 350
 DB 257 KESRLEDTKHRVDFLQMLDMSQNSKXETESHKALSDELVAQSIIIFAGYETTSVLSF 316

```

351 MFHRLSEDKAVQDKLREEICQ1--DTDMPTLDLNLPLYEAFVKESLURLDPPSPYANRE 408
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 IMYELATHPDVQOKLQBEIDAVLPNKAPPTYDTVLQMEYLDWVNVETLRLFPILMRERV 376
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
409 CLKDEDFIPLAEPVIGRDGSGVINEVRIITKGTVMVLPFLNINRSKFIYGEDAEETPRRWL 468
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
377 CKKQVE-----INGMTPKGVVMIPSAVLRDP-KYWTBEKFLPERFS 420
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
469 EDVTDLSLS-IEAPVGHQAGFISGPACFGWRFAVAEMKAFILVTLRAVQFEP 520
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
421 KXNKNDIDPIYTFFG-----SGPRNCIGRPALNMKALIRVLQNFSEKP 467
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
AAR81464
ID AAR81464 standard; protein; 503 AA.
XX
XX AAR81464;
AC
XX
XX 01-AUG-1996 (first entry)
XX
XX Human derived cytochrome P4503A4.
XX
XX Human derived cytochrome; P4503A4; commercial cDNA library; Yeast;
XX transfection; recombinant production; expression vector; mammal;
XX immunisation; sensitisation; antibody; determination; detection;
XX non-cross reactive.
XX
XX Homo sapiens.
OS
XX
XX JF08027197-A.
XX
XX 30-JAN-1996.
XX
XX 13-JUL-1994; 94JP-00161552.
XX
XX 13-JUL-1994; 94JP-00161552.
XX
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX WPI; 1996-136338/14.
XX
XX N-PSDB; AAT17399.
XX
XX Antibody recognising human derived cytochrome P4502A4 - allows specific
XX detection of cytochrome P450 species in humans.
XX
XX Example 1; Page 10-12; 13pp; Japanese.
XX
XX The present sequence is the human derived cytochrome (HDC) P4503A4, which
XX was obtd. from a commercial cDNA library. Yeast were transfected with an
XX expression vector contg. the HDC cDNA, cultured and then disrupted to
XX give a microsomal fraction. The HDC was purified from the fraction, and
XX used to immunise and sensitise a mammal. Blood was drawn from the mammal,
XX and an anti-HDC antibody isolated. The antibody obtd. recognises HDC
XX P4503A4, partic. at a serum dilution rate of 1:10000. and is
XX substantially without cross reaction to other HDC P450 spp
XX
XX Sequence 503 AA;
XX
XX Query Match 13.9%; Score 396.5; DB 2; Length 503;
XX Best Local Similarity 28.7%; Pred No. 1.3e-25;
XX Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24
XX
XX 13 LAAFSW--ASIAFFSLYIAPRRS-SLYNLQG-PNHNTY-FTGNFIDILSARTG-----E 61
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 7 LAMETWLLLAVALSVILYLYGTHSHGLFKLIGPGTLPFLGN---ILSYHKGFQCFMDME 63
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 62 EHAKYREKYGSTLRFAGIAGAPVLNSTDPKVENHW-KEAYDY---PKPGMAARVLRIAT 117
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 64 CHKKYKGVNGF---YDG---QQPVLAITDPMIKVLVKECYSVFTNRFRPGVGFMKSA- 117
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 118 GDGVVTAEGEAHKHRRIRIMPSLSAQAVKSMVP;FLEKGMELVDKMWDEDEKDMVAGES 177
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	118	---	ISIADEEWEKLRSLSPFTTSGKJEMVPIIAQYGDVLRNLRREA-----	164
Qy	178	AGEKATRLT-	EGVDYKDWVGRATLDVWALAGFDYKSDLOQKTNELVAFVGLTDGFA	236
Db	165	-----	ETGKPVTLKDVFGAYSDVITSTSGVNISSLNPNQD-----	201
Qy	237	PTLSFKAIM-	WDVF-PYRTMK-RRHEIPLTOGLAV---SRRVGIELMEQKQAVLGSA	290
Db	202	PFVENTKLLRFDLDFDFFLSITVFPFPILEVLNLCVFPREVTFLEKSKVKR	---	256
Qy	291	SDQAVDKDVQGRDILSLVLRAMIANLPESQKLSDEEVLQAIQSNLLFAGYETSTVLT	350	
Db	257	KESRUEDTQKURVDFLQMLDMSQNSKETESHKALSDELVAQSIIFIPAGYETSSVLSF	316	
Qy	351	MFHRLSEDKAVQDKLRBEICQI--	DTDMPTLDEINALPYLEAFVKESLRDLPSPYANRE	408
Db	317	IMVELATHPDVQKQLEBEIDAVLENKAPPTYDTVLQMEYLDVMVNETLRFPIAMRLERV	376	
Qy	409	CLKDEDFIPLAEPVIGRDGVSINERVITKGTVMWMLPLFNINRSKFYIGSDAEFRPERWL	468	
Db	377	CKKQVE-----	INGNFIPLKGVVNIIPSYALHRDP-KYWTPEKFLPERFS	420
Qy	469	EDVTDLSLNS-IEAPYGHQASFI	SGPRACFGWRFAVAMKAFLEVTLLRVOFBP	520
Db	421	KKNKONIDPIYVTPFG-----	SGPRNCIGWRPALMMKALIRVLQNFSPKP	467
RESULT 7				
AAR93170				
ID	AAR93170	standard; protein; 503 AA.		
AC	AAR93170;			
CC				
XX				
DT	11-OCT-1996	(first entry)		
DE				
XX				
XX				
KW	Human cytochrome P450 molecular species 3A4 protein.			
KW	Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;			
KW	liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;			
KW	evaluation; safety; fusion protein; metabolite; detoxification;			
KW	carcinogenic.			
XX				
OS	Homo sapiens.			
XX				
FN	JP08056695-A.			
XX				
PD	05-MAR-1996.			
XX				
FF	15-JUL-1994; 94JP-00164184.			
XX				
PR	20-JUL-1993; 93JP-00201120.			
PR	30-JUL-1993; 93JP-00208279.			
PR	17-JUN-1994; 94JP-00136053.			
XX				
PA	(SUMO) SUMITOMO CHEM CO LTD.			
XX				
DR	WPT; 1996-182311/19.			
DR	N-PSDB; AAT28383.			
XX				
PT	Novel method for the evaluation of the safety of a cpd. - using a human			
PT	cytochrome P450 and yeast NADPH reductase to determine whether the			
PT	analyte cpd. is detoxified or metabolised to a carcinogen.			
XX				
PS	Example 1; Page 24-26; 74pp; Japanese.			
XX				
CC	This is the amino acid sequence of the human cytochrome P450 molecular			
CC	species 3A4 protein. The corresp. gene was amplified from a human liver			
CC	derived cDNA library as 2 fragments of 0.6 and 0.9 kb using primers			
CC	AAT26933-6. The prod. was cloned into the yeast expression vector pAAH5N			
CC	to generate plasmid p3A4 for prodn. of the cytochrome only or into the			
CC	vector pAHR to generate the plasmid p3A4R for co-prodn. with the yeast			
CC	NADPH-P450 reductase. The sequence is placed under control of the yeast			

CC ADH gene promoter and terminator. The vectors are used in a method for
CC evaluating the safety of a cpd. by reacting the test cpd. with
CC recombinantly produced human cytochrome P450 mol. species 1A2 (AA28380),
CC 2C9 (AA28381), 2E1 (AA28382), 3A4 or their variants (AA28384-98)
CC together with yeast NADPH-P450 reductase (either as a fused protein or as
CC a cell extract) and analysing the resultant metabolite. The cpd. is
CC considered "safe" if it is detoxified or not rendered carcinogenic or
CC "unsafe" if it is not detoxified or is metabolised to a carcinogenic cpd
XX
SQ Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;
Best Local Similarity 28.7%; Pred. No. 1.3e-25;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLLAVALSVLLVLYGTHSHGLFKLGIPOFTPLFLGN---ILSVHKGFCEFDME 63
QY 62 EHAKEYREKSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYDY---PKPGMAARVLRIAT 117
DB 64 CHKYGKVGWF---YDG--QQPVLAITDPDMIKTVLKECVSVFTNRRPFGVGFNKSA- 117
QY 118 GDGVVTAEGEAHRRIMIPSLSAQAVKSMVPFLEKGMELVDKMDAAEKDMAGVES 177
DB 118 ---ISIAEDEEWKRLSLLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164
QY 178 AGEKATRLT-EGVDVKDWGRATLDYMALAGFDYKSDSLQNTNELYVAFVGLTDGFA 236
DB 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNIIDSLNPDQ----- 201
QY 237 PTLDSFKAIM-WDFV-PYFRMTK--RRHEIPTQLAV---SRRVGIELMBQKQAVLGS 290
DB 202 PFVENTKLLRFDLDPDFLSITVPFPILEVLNICVFPREVTFNLRKSVKR-----M 256
QY 291 SDQAVKDKVQGRDILSLVRANIAANLPESQKLSDEEVLQAISNLLFAGYETSSVLTW 350
DB 257 KESLEDTQKRVDFLQLMIDSONSKETESHKALSDELVAQSIIFIFAGYETSSVLSF 316
QY 351 MFHRLSEDKAVQDKLREICQI--DTDMPTLDELNALPYLEAFVKESLRDLPSPYANRE 408
DB 317 IMVELATHPDVQKQLEEDIDAVLPNKAPPTVDYTLQMEYLDVMVNETLRFPFIAMRLERV 376
QY 409 CLKDEDFIPLAEFVIGRDSGVINEVRITKGTVMVLPFNINRSKFIFYGDAEAEFRPERWL 468
DB 377 CKKQDV-----INGMFIPKGVVWVIMPSYALHSDP--KYWTEPEKELPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFGSPRACFGWRFVAEMKAFVLTLRVQFEP 520
DB 421 KKKKNIDPIYTPFG-----SGPRNCIGMRFALMNKKLALIRVLQNFSEKP 467

RESULT 8
AA05202 ID AA05202 standard; protein; 503 AA.
XX AC AA05202;
XX DT 17-JUN-1999 (first entry)
XX DE Human CYP3A4 protein.
XX KW CYP3A4 gene polymorphism; polymorphic locus; human; altered metabolism;
XX KW CYP3A4 substrate; drug-drug interaction identification; toxin exposure;
XX KW genetic linkage detection; phenotypic variation.
XX OS Homo sapiens.
XX FN W09913106-A1.
XX PD 18-MAR-1999.
XX PF 02-SEP-1998; 98WO-US018158.

XX 10-SEP-1997; 97US-0058612P.
XX (AXYS-) AXYS PHARM INC.
XX PI Lichter JB, Guida M;
XX WPI; 1999-215070/18.
XX N-PSDB; AAX28295.
XX New isolated CYP3A4 polymorphic sequences.
XX Disclosure; Page 25-27; 40pp; English.
CC This sequence represents the human CYP3A4 gene protein. The invention
CC relates to a CYP3A4 sequence polymorphism, which is part of a non-
CC naturally occurring chromosome. Nucleic acids comprising the CYP3A4
CC polymorphic sequences can be used to screen patients for altered
CC metabolism for CYP3A4 substrates, potential drug-drug interactions, and
CC adverse/side effects as well as diseases that result from environmental
CC or occupational exposure to toxins. They can also be used to establish
CC animal, cell culture and in vitro cell-free models for drug metabolism.
CC Polymorphic CYP3A4 gene sequences can be used for expression studies to
CC determine the effect of promoter and/or intron sequence variations on
CC mRNA expression and stability. The polymorphisms are also used as single
CC nucleotide polymorphisms to detect genetic linkage to phenotypic
CC variation in activity and expression of CYP3A4. The nucleic acids can
CC also be used to generate genetically modified non-human animals or site
CC specific gene modifications in cell lines
XX
SQ Sequence 503 AA;

Query Match 13.9%; Score 396.5; DB 2; Length 503;
Best Local Similarity 28.7%; Pred. No. 1.3e-25;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLLAVALSVLLVLYGTHSHGLFKLGIPOFTPLFLGN---ILSVHKGFCEFDME 63
QY 62 EHAKEYREKSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYDY---PKPGMAARVLRIAT 117
DB 64 CHKYGKVGWF---YDG--QQPVLAITDPDMIKTVLKECVSVFTNRRPFGVGFNKSA- 117
QY 118 GDGVVTAEGEAHRRIMIPSLSAQAVKSMVPFLEKGMELVDKMDAAEKDMAGVES 177
DB 118 ---ISIAEDEEWKRLSLLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164
QY 178 AGEKATRLT-EGVDVKDWGRATLDYMALAGFDYKSDSLQNTNELYVAFVGLTDGFA 236
DB 165 -----ETGKPVTLKDVFGAYSMDVITSTSGVNIIDSLNPDQ----- 201
QY 237 PTLDSFKAIM-WDFV-PYFRMTK--RRHEIPTQLAV---SRRVGIELMBQKQAVLGS 290
DB 202 PFVENTKLLRFDLDPDFLSITVPFPILEVLNICVFPREVTFNLRKSVKR-----M 256
QY 291 SDQAVKDKVQGRDILSLVRANIAANLPESQKLSDEEVLQAISNLLFAGYETSSVLTW 350
DB 257 KESLEDTQKRVDFLQLMIDSONSKETESHKALSDELVAQSIIFIFAGYETSSVLSF 316
QY 351 MFHRLSEDKAVQDKLREICQI--DTDMPTLDELNALPYLEAFVKESLRDLPSPYANRE 408
DB 317 IMVELATHPDVQKQLEEDIDAVLPNKAPPTVDYTLQMEYLDVMVNETLRFPFIAMRLERV 376
QY 409 CLKDEDFIPLAEFVIGRDSGVINEVRITKGTVMVLPFNINRSKFIFYGDAEAEFRPERWL 468
DB 377 CKKQDV-----INGMFIPKGVVWVIMPSYALHSDP--KYWTEPEKELPERFS 420
QY 469 EDVTDLSNS-IEAPYGHQASFGSPRACFGWRFVAEMKAFVLTLRVQFEP 520
DB 421 KKKKNIDPIYTPFG-----SGPRNCIGMRFALMNKKLALIRVLQNFSEKP 467

Matches 146; Conservative 96; Mismatches 218; Indels 133; Gaps 25;

QY 7 LTGALGLAPSW--ASIAFSLYLAPRRSSLYNLOGPNHTNY--FTGNFLDILSARTGE 62
 DB 17 LAGAVASLLMLVAVTLEWAWTWRDLRAQGLNGTRYLFTGDLRE--TARVNR 74
 QY 63 -----HAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNNHVMKEAYD 102
 DB 75 ARKKPLPLGCHDITPRVQPMHSTIKYK-LSFTMGPTFRVMIPTDELKVELSNKFG 133
 QY 103 Y----PKPGMAARVLRTATGCVVTAAGEAKHRRIMIPSLSAQAKSMWVIFLEKGMEL 159
 DB 134 HFGKPRSSRIGRL-----ANGLVNDHGWAKHRRILNAPAFHHEKIKGMVFTCCIE 189
 QY 160 VDKWMDAAEKMAVGESAGEKKATRLTEGVDVXDWVGRATLDVWALAGF--DYKSD-- 215
 DB 190 ITRW-----DNSMS-SEGSE-----IDVWPEFQNLTDGVIISRTAFGSNYQSGRR 233
 QY 216 --SLQNTKNELYVAVGLTDGFAPTLDSPKATMWFVP--YPRTKRHRHEIPTLQGLAVS 271
 DB 234 IFELQELAEERLT-----QSVQTIPIFYFWFLPTKNNRM----- 268
 QY 272 RRVGIELMECKQAVLGASDQAVDKKDYQGR-----DILSLVLRANT--AANLPESOKLS 325
 DB 269 RAIDVEIRKILRE-IIGKR-----EKDTKRETNKDLLGLLESNTROSGNASLGLT 321
 QY 326 DEVLQAIQNLNLPAGYETSTVLTMFHLRLSDKAVQDKLREI--CQIDTMDPTLDELNA 384
 DB 322 TEDVIECKLFYFAGMETTSVLLTWLTVLSMHPEWQERAREVLSHFGRTPDYDSLRS 381
 QY 385 LPYLEAFVKESSLRLDPPSPYANRECKDEDFIPLAEPVIGRDSVINEVRITKGTWMLP 444
 DB 382 LKTVTWILHEVLRYPPATFLTRTYKEME-----LGGIKTPAGVELLLP 426
 QY 445 LFNINRSKFIYGEDAEFRPERWLEDVTDLSLNSIEAPYGHQASFI---SGPRACFGWRFA 501
 DB 427 VIFIHDPDIDGKQAGEFNPERP-----ANGISSATRHQAAPFPFGGPRICIGQSFA 479
 QY 502 VAEMKAFVTLRVOFE--PIISHPEYSHITLISRPVIGREKEXQMBLQ 552
 DB 480 LLEAKMTLCTILRFGSELSPSYTHAPYTVITL-----HPQHGAIKRLK 523

RESULT 11
 AAEE22852
 ID AAEE22852 standard; protein; 503 AA.
 XX
 XX AAEE22852;
 DT 09-AUG-2002 (first entry)
 XX
 DE Human cytochrome P-4503A7 (CYP3A7) protein.
 XX
 KW Human; cytochrome P-4503A7; CYP3A7; promoter; environment pollutant;
 KW digital nuclear receptor binding motif; drug bioavailability; CAR;
 KW pregnane activated receptor; PAR; constitutive androstane receptor;
 KW drug interaction; toxicity.
 XX
 OS Homo sapiens.
 XX
 XX WO200224918-A1.
 XX
 XX 28-MAR-2002.
 XX
 XX 19-SEP-2001; 2001WO-SE002007.
 XX
 XX 22-SEP-2000; 2000SE-00003393.
 XX
 XX (BIOV-) BIOVITRUM AB.
 XX
 XX Berkenstam A, Bertilsson G, Blomquist P;
 XX WPI; 2002-394140/42.
 XX

DR N-PSDB; AAD36214.
 XX
 PT New isolated human cytochrome P-450347 promoter region useful in
 PT screening for pharmacological agents comprises digital nuclear receptor
 PT binding motifs.
 XX
 PS Disclosure; Page 34-36; 39pp; English.
 XX
 CC The invention relates to human cytochrome P-4503A7 (CYP3A7) promoter
 CC comprising digital nuclear receptor binding motifs (dnrs). CYP3A7
 CC promoter sequences are useful for identifying agents that modulate their
 CC activity. They are also useful for identifying agents that modulate the
 CC association of CYP3A7 promoter and transcription factor e.g. Pregnane
 CC Activated Receptor (PAR) or Constitutive Androstane Receptor (CAR). The
 CC agents identified by the CYP3A7 promoter are useful in modulating
 CC metabolism of endogenous and/or exogenous compounds, drug interaction,
 CC toxicity and/or bioavailability of drugs. The CYP3A7 is also useful for
 CC screening environment pollutants for potential endocrine disrupting
 CC capacity. The present sequence is human CYP3A7 protein
 XX
 SQ Sequence 503 AA;
 Query Match 13.5%; Score 385.5; DB 5; Length 503;
 Best Local Similarity 29.0%; Pred. No. 1.2e-24;
 Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

QY 13 LAAPSW--ASTAFSLYL--APRRSSLYNLOG--PNHTNY--FTGNFLDILSARTG-----E 61
 DB 7 LAVETWLLAVALLLYLGRTRTHGLFKLGIPTPLPLGN---ALSFRKGTWTFDME 63
 QY 62 BHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNNHVM--KEAYD-----YKPGMAARVL 113
 DB 64 CYKTKRWG-----IYDCQCPMLAITDPDMIKTLVKECYSVFTNRRPFGVGMKNAI 118
 QY 114 RIATGCVVTAAGEAKHRRIMIPSLSAQAKSMWVIFLEKGMELVDKMDAAEKDMA 173
 DB 119 SI-----AEDEEWKIRSLSPFTTSGKLKEMVPIIAQYGDVLVRNLREA----- 164
 QY 174 VGESAGEKKATRLT--EGVDVXDWVGRATLDVWALAGFDYKSDSLQNTKNELYVAVGLT 232
 DB 165 -----ETGKPVTLKHVFGAYSMDVITSTSGVSIDSLSNPNQD---PFVENT 207
 QY 233 D---GFAPTLDSFKAIMWDFVYPRTKRHRHEIPTLQGLAVS---RRVGIELMECKQAV 286
 DB 208 KKLLRFNP--LDPFVLSIKVF--PFL-----TPIELALNITVPRKVISFLTKSVKQIK 257
 QY 287 LGSASDAQVKKDYQGR--DILSLVLRANIAANLPESOKLSDEVLQAIQNLNLPAGYETS 344
 DB 258 EGRLL-----KETQKHRVDFQLMIDSONSKDSETHKALSDLELMAQSIIFIFAGYETT 310
 QY 345 STVLTMFHLRLSEDKAVQDKLREIQCIDTDM-----PTLDELNALPYLEAFVKESURLD 399
 DB 311 SSVLSFIIVELATHDPVQKQVKE---IDTVLPNKAPPTVDTVLQLEYLDWVWNETLRLF 367
 QY 400 PSPVYANRECKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDA 459
 DB 368 PYAMLERVCKKQV-----INGMFIPIKGVVWMIPIPSVYLHDP--KYWTPE 411
 QY 460 EBFPRERMELEDVTDLSNS--IEAPYGHQASFISSGPRACFGWRFAVEMKAFVTLRVOF 518
 DB 412 EXFLPERFSKKNKONIDPVIYTPFG-----SGPRNCIGMRFALVNMKALVRLQNF 465
 QY 519 EP 520
 DB 466 KP 467

RESULT 12
 ABG68747
 ID ABG68747 standard; protein; 535 AA.
 XX
 XX ABG68747;
 XX

DT 07-OCT-2002 (first entry)
DE Cytochrome P450 CYP3A7 and CYP3A1 fusion protein.
XX
XX
KW Cytochrome P450; CYP3A1; CYP3A2; CYP3A4; CYP3A5; CYP3A7;
KW drug metabolism; drug design; drug screening.
XX
XX Homo sapiens.
OS
XX WO200244213-A1.
PN
XX
XX 06-JUN-2002.
PD
XX
XX 28-NOV-2001; 2001WO-SE002631.
PF
XX
XX 28-NOV-2000; 2000SE-00004366.
PR
XX 11-JUN-2001; 2001SE-00002061.
PR
XX (ZAPH/) ZAPHIROPOULOS P G.
PA (FINT/) FINTA C.
FA
XX
XX Zaphiropoulos PG, Finta C;
Pi
XX
XX WPI; 2002-557532/59.
DR N-PSDB; ABK97643.
DR
XX
XX Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
PT metabolism, in drug design and drug screening.
XX
XX
XX Claim 1; Page 73-75; 131pp; English.
PS
XX
XX The invention describes a cytochrome P450 protein (I) in which CYP3A43
CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
CC fragments, variants and multiples of (I) having essentially the same
CC characteristics (I) is useful as a medicament, and for evaluating drug
CC metabolism, in drug design, and drug screening, and in tests for
CC adjusting the dose of drugs. This is the amino acid sequence of a novel
CC cytochrome P450 protein
XX
XX Sequence 535 AA;
SQ

Query Match 13.5%; Score 385.5; DB 5; Length 535;
Best Local Similarity 29.0%; Pred. No. 1.3e-24;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61
DB 7 LAVETWLLAVSLILLVLYGTRTHGLFKLGIPOGTELPFLGN---ALSFRKGYWTFDME 63
QY 62 BHAKYREKYGSTLRFAGIAGAPVLSNDPKVFNHVM-KEAYD-----YPRKGMARVL 113
DB 64 CYKRYKRWG-----YDCCQPMIAITDPMIKTVLKECYSVFTNRPPGPGVGFKNKAI 118
QY 114 RIATGDGVVTAEGEAKHRRIMIPSLSAQAVKSWVPFLEKGMELVDKQMDAEKDMA 173
DB 119 SI-----ADEENKGIKRLSLLFTTSGLKKEWPIIAQYGDVLVRLREA----- 164
QY 174 VGESAGEKKATRLT-EGVDYKDWGVRATLDVMAAGFDYKSDLSQNKTNELVAVGLT 232
DB 165 -----ETGKEVTLKHFVFGAYSDVITSTSGVSDLSLNNPDQ-----PVEVNT 207
QY 233 D---GFAPTLDSFKAIMNDVFPVYFRTMKRRHEIPLTQGLAVS---RRVGELEMEQKQAV 286
DB 208 KXLLRFPN-LDPFVLISIKVF-PFL-----TILEALNITVPFKVISFLKSVKQIK 257
QY 287 LGSASDAQVKKQVGR--DILSLVRANIANLPESQKLSDEVLQAQINLLPAGYETS 344
DB 258 EGRL-----KETQKRVDFLQMLDMSQSKDSETHKALSDELMAQSIIFAGYETT 310
QY 345 STVLTMFHRLESKAVODKLRKEICQIDTDM-----PTLDELNALPYLEAFVKESIRLD 399
DB 311 SSVLSFIITELATHPDVQOKVQKE---IDTVLPKAPPTDVLQLELYDMVNVNELRLF 367

QY 400 PPSPYANRECLKDEDFIPLAEPVIGROGSVINEVITKGTVMPLPLFNINRSKFIYGEDA 459
DB 368 FVAMLERVCKKQVE-----INGMFIKGVVVMPISYVLHDP-KYWTEP 411
QY 460 EBFRRPERWLEDVTDLSLNS-IEAPYGHQASFIISGRACFGWFAVAEKAFILVTLRRVQF 518
DB 412 EKFLPERFSKKNKNDIPVIYTPFG-----SGPRNCITGMRFALVNMKMLALVRLQNF 465
QY 519 EP 520
DB 466 KP 467

RESULT 13
AAE26192
ID AAE26192 standard; protein; 502 AA.
XX
AC AAE26192;
XX
DT 14-NOV-2002 (first entry)
XX
XX Human cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) protein.
DE
XX
XX Human; cytochrome P450; subfamily IIIA; polypeptide 5 isogene; CYP3A5;
KW drug screening; polymorphism; haplotype; drug metabolising disorder;
XX gene therapy.
XX
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FT Misc-difference 30 /note= "This amino acid changes to Tyr due to single
FT nucleotide polymorphism"
FT Misc-difference 100 /note= "This amino acid changes to Tyr due to single
FT nucleotide polymorphism"
XX
XX WO200246209-A2.
XX
XX 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US047218.
XX
XX 08-DEC-2000; 2000US-0254367P.
PR 03-MAY-2001; 2001US-0268470P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Han J, Klieem SE, Rounds E;
PI
XX WPI: 2002-636448/68.
DR N-PSDB; AAD43350, AAD43351.
XX
XX Novel isolated polynucleotide which is a polymorphic variant of
PT cytochrome P450, subfamily IIIA, polypeptide 5 (CYP3A5) gene useful for
PT expressing CYP3A5 protein isoform used in drug screening techniques.
XX
XX Claim 29; Fig 3; 127pp; English.
PS
XX
XX The invention relates to isolated polynucleotide having cytochrome P450,
CC subfamily IIIA, polypeptide 5 isogene (CYP3A5). The invention is useful
CC for screening drugs. The invention is useful for studying expression and
CC function of CYP3A5 and expressing CYP3A5 protein for use in screening for
CC candidate drugs to treat diseases related to CYP3A5 activity. The
CC polymorphism and haplotype data is useful for validating whether CYP3A5
CC is a suitable target for drugs to treat drug metabolising disorders,
CC screening for such drugs and reducing bias in clinical trials of such
CC drugs. The invention is also useful for therapeutic purposes. The
CC invention is useful in studying the effect of variation on the biological
CC activity of CYP3A5 as well as on the binding affinity of candidate drugs
CC to CYP3A5, or for studying the enzymatic properties of such CYP3A5
CC variants using these candidate drugs as substrate. The invention is

QY 405 ANRECLXDEDFIPLAEVIGRDSVINEVRITKGTWVWMLPLFNINRSKFIYGEDAEFRP 464
DB 373 LERTCKXKDE-----INGVIFPGSMVVIPTVALHDP-KYWTBEEFRP 416
QY 465 ERWLEDVTDLSINS-IEAPYGHQASFGSPACFGWFAVAEMKAFUFTVLRVQPEP 520
DB 417 ERFSKK-KDSIDPIYTPFG-----TGPRNCIGMRFALMNMKALIRVLQNFSPKP 466

RESULT 15
ADE57190
ID ADE57190 standard; protein; 502 AA.
AC ADE57190;
DT 29-JAN-2004 (first entry)
XX Human Protein P20815, SEQ ID NO 3051.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
OS WO2003016475-A2.
PN 27-FEB-2003.
PD 14-AUG-2002; 2002WO-US025765.
PF 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
PI Woolf C, D'urso D, Befort K, Costigan M;
DR WPI; 2003-268312/26.
DR GENBANK; P20815.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 502 AA;
Query Match 13.5%; Score 385; DB 7; Length 502;
Best Local Similarity 28.1%; Pred. No. 1.3e-24;
Matches 151; Conservative 84; Mismatches 196; Indels 106; Gaps 25;
QY 13 LAAPSW--ASIAFFSLYL-APRRSSLYNLOG-PNHTNY-FTGNFLDILSARTG-----E 61
DB 7 LAVETWLLAVSLVLYLYGTGTHGLFKRLGPGTPLPLGN-----VLSYQGLWKPFTE 63
QY 62 EHAKYREKYGSTLRFPAGIAGAPVLNSTDPKVFNHYM-KEAYDYPKPGNAARVLRATG-- 118
DB 64 CYKYGKMWGT---YEG--QLPVLAITDPDVRTVLVKECY-----SVFTNRSLSGPV 111
QY 119 ----DGVVTAEGEAHRRIRIMIPSLSAQAQKSMVPIFEKGMELVYKMMEDAAEKDVA 174
DB 112 GFMKSAISLAEDENKRIIRSLSPFTTSGKLKXEMFPIIAQYGDVILVRLNRR-ABKG-- 167
QY 175 GESAGEKKAIRLETEGVGVKDMVGRATLDVMALAGFDYKSDSLQN-----KTNELYVAF 228
DB 168 -----KPVTLKDIFGAYSMVDVITGTFGVNIDSLNPNQDPFVESTKKELKF 213
QY 229 VGLTDGFPAPTLDSFKAIMWDFVPIFTMKRHEIPLTQGLAVS--RRVGIELMEQKQAV 286
DB 214 ----GFLDPL-FLSIILPPFL-----TPVFEALNVSLFPKDTINFLSKS---- 252
QY 287 LGSASDAQVDKDVQGRDILSLVRAIANLNPESOKLSDEEVLAQISNLLPAGVETST 346
DB 253 VNRMKSLNDKOKHRLDFLQLMIDSONKETESHKALSDELAQAQSIIFIPAGVETSS 312
QY 347 VLTWMFRLSEDKAVQDKLREEICQI--DTDMPTDELNALPYLEAFVKESIRLPPSPY 404
DB 313 VLSPTLYELATHPDVQOKLQKEIDAVLPNKAPPTVDVVQMEYLDVMVNETLRLFPVAIR 372
QY 405 ANRECLXDEDFIPLAEVIGRDSVINEVRITKGTWVWMLPLFNINRSKFIYGEDAEFRP 464
DB 373 LERTCKXKDE-----INGVIFPGSMVVIPTVALHDP-KYWTBEEFRP 416
QY 465 ERWLEDVTDLSINS-IEAPYGHQASFGSPACFGWFAVAEMKAFUFTVLRVQPEP 520
DB 417 ERFSKK-KDSIDPIYTPFG-----TGPRNCIGMRFALMNMKALIRVLQNFSPKP 466

Search completed: April 2, 2004, 13:58:37
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17.5 Seconds
(without alignments)
1643.181 Million cell updates/sec

Title: US-10-066-007-3
Perfect score: 2852
Sequence: 1 MFILVLTGALGAARSWAS.....RIVGREKGYQMRLOVKPVE 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	4	US-09-518-386B-1
2	2852	100.0	557	4	US-09-518-386B-3
3	396.5	13.9	503	4	US-09-144-357-2
4	370	13.0	504	1	US-08-457-274A-25
5	370	13.0	504	5	PCT-US95-05758-25
6	367	12.9	520	4	US-09-527-073-2
7	345	12.1	524	4	US-09-976-594-533
8	338.5	11.9	503	4	US-09-583-447A-2
9	331	11.6	540	4	US-09-302-620B-99
10	330.5	11.6	507	1	US-08-457-274A-22
11	330.5	11.6	507	5	PCT-US95-05758-22
12	328	11.5	504	4	US-09-583-447A-4
13	328	11.5	540	4	US-09-302-620B-98
14	326.5	11.4	541	3	US-09-158-767-19
15	326.5	11.4	541	3	US-09-158-767-20
16	321	11.3	576	3	US-08-948-564-16
17	318.5	11.2	526	1	US-08-298-426-4
18	294.5	10.3	489	4	US-09-852-087-4
19	294	10.3	522	4	US-09-302-620B-97
20	289	10.1	522	4	US-09-302-620B-96
21	276	9.7	420	4	US-09-583-447A-6
22	276	9.7	467	4	US-09-126-420A-17
23	269.5	9.4	517	4	US-09-302-620B-100
24	269.5	9.4	517	4	US-09-911-781-32
25	269	9.4	507	1	US-08-457-274A-23
26	269	9.4	507	5	PCT-US95-05758-23
27	267.5	9.4	517	4	US-09-302-620B-101

28	263	9.2	523	4	US-09-302-620B-95
29	259.5	9.1	510	3	US-08-948-564-4
30	259	9.1	510	4	US-09-852-067-2
31	251.5	8.8	508	4	US-09-126-420A-25
32	246	8.6	512	4	US-09-302-620B-103
33	243	8.5	512	4	US-09-302-620B-102
34	241.5	8.5	476	1	US-08-313-075A-30
35	236.5	8.3	498	1	US-08-457-274A-24
36	236.5	8.3	498	5	PCT-US95-05758-24
37	235.5	8.3	499	4	US-09-302-620B-104
38	235.5	8.3	504	4	US-09-126-420A-18
39	233.5	8.2	509	4	US-09-499-302A-8
40	228	8.0	426	2	US-08-560-398-4
41	227	8.0	504	4	US-09-499-302A-6
42	226.5	7.9	500	3	US-09-292-768-68
43	226.5	7.9	500	3	US-09-292-768-70
44	225.5	7.9	500	3	US-09-292-768-4
45	223.5	7.8	524	4	US-09-126-420A-24

ALIGNMENTS

RESULT 1
US-09-518-386B-1
; Sequence 1, Application US/09518386B
; Patent No. 6365386
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518,386B
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: TRANSIT
; LOCATION: (1)..(26)
US-09-518-386B-1

Query Match	100.0%;	Score 2852;	DB 4;	Length 557;
Best Local Similarity	100.0%;	Pred. No. 4.4e-278;	Mismatches 0;	Indels 0;
Matches 557;	Conservative 0;			Gaps 0;
QY	1	MFILVLTGALGAARSWASIAFFSLYLA	PRSSLYN	LOGPNHTNYFTGNFLDILSARTG 60
Db	1	MFILVLTGALGAARSWASIAFFSLYLA	PRSSLYN	LOGPNHTNYFTGNFLDILSARTG 60
QY	61	BEHAKYREKYGSTLR	PAGIAGAVLN	STDPKVFNVHVKAYDYPKPGMAARVLRIATG 120
Db	61	BEHAKYREKYGSTLR	PAGIAGAVLN	STDPKVFNVHVKAYDYPKPGMAARVLRIATG 120
QY	121	VVTEGAHKKRRRIMIPSL	SAQAVKSWVP	IFLEKGMELVDKMWEDAAEKDMAVGESAGE 180
Db	121	VVTEGAHKKRRRIMIPSL	SAQAVKSWVP	IFLEKGMELVDKMWEDAAEKDMAVGESAGE 180
QY	181	KKATRLTEGVVDKDWVGRATLD	VMALAGFDYKSD	SLQNKTNELYYAVFVLGTGDFAPTLD 240
Db	181	KKATRLTEGVVDKDWVGRATLD	VMALAGFDYKSD	SLQNKTNELYYAVFVLGTGDFAPTLD 240
QY	241	SFKAIMDVPYFRTMKRRHEI	PLTQGLAVSR	RVGIELMEQKQAVLGSASQAVDKDV 300
Db	241	SFKAIMDVPYFRTMKRRHEI	PLTQGLAVSR	RVGIELMEQKQAVLGSASQAVDKDV 300

QY 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWPHRLSEDKA 360
 DB 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWPHRLSEDKA 360
 QY 361 VODKLRBEICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAE 420
 DB 361 VODKLRBEICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAE 420
 QY 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFEPRERWLEDVTDLSNLSIEA 480
 DB 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFEPRERWLEDVTDLSNLSIEA 480
 QY 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540
 DB 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540
 QY 541 GREKEGYQMRLOVKPVE 557
 DB 541 GREKEGYQMRLOVKPVE 557

RESULT 2
 US-09-518-386B-3
 ; Sequence 3, Application US/09518386B
 ; Patent No. 6365386
 ; GENERAL INFORMATION:
 ; APPLICANT: HOSHINO, Tatsuo
 ; APPLICANT: OJIMA, Kazuyuki
 ; APPLICANT: SETOGUCHI, Yutaka
 ; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
 ; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
 ; CURRENT APPLICATION NUMBER: US/09/518,386B
 ; CURRENT FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: EP 99104668.1
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: EP 00101666.6
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 557
 ; TYPE: PRT
 ; ORGANISM: Phaffia rhodozyma
 US-09-518-386B-3

Query Match 100.0%; Score 2852; DB 4; Length 557;
 Best Local Similarity 100.0%; Pred. No. 4.e-278;
 Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFILVLLTGALGLAASFWSIASIAFFSLYLAPRSSLYNLOGPNHTNYFTGNFLDILSARTG 60
 DB 1 MFILVLLTGALGLAASFWSIASIAFFSLYLAPRSSLYNLOGPNHTNYFTGNFLDILSARTG 60
 QY 61 EEHAKYREKYGSTLRAGTAGAPVLNSTDPKVFNHVMKEAYDYPKPGMAARVLRATGDG 120
 DB 61 EEHAKYREKYGSTLRAGTAGAPVLNSTDPKVFNHVMKEAYDYPKPGMAARVLRATGDG 120
 QY 121 VVTAEGSAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKOMAVGESAGE 180
 DB 121 VVTAEGSAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKOMAVGESAGE 180
 QY 181 KKAATRLTEGVDDKDWVGRATLDVMAAGFDYKSDSLQNKTNELVYVAVGLTDGFPATLD 240
 DB 181 KKAATRLTEGVDDKDWVGRATLDVMAAGFDYKSDSLQNKTNELVYVAVGLTDGFPATLD 240
 QY 241 SFKAIMWDFVYFRTMKRREIIFLTQGLAVSRVAGIELMEQKQAVLGSASQAVDKDV 300
 DB 241 SFKAIMWDFVYFRTMKRREIIFLTQGLAVSRVAGIELMEQKQAVLGSASQAVDKDV 300
 QY 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWPHRLSEDKA 360
 DB 301 QGRDILSLVRANIANLPESOKLSDBEVLAIQISNLLFAGYETSTVLTWPHRLSEDKA 360

QY 361 VODKLRBEICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAE 420
 DB 361 VODKLRBEICQIDTDMPTLDELNALPYLEAFVKESLRDPPSPYANRECKDEDFIPLAE 420
 QY 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFEPRERWLEDVTDLSNLSIEA 480
 DB 421 PVIGRDGSVINEVRITKGTWMLPLFNINRSKFIYGEDAEFEPRERWLEDVTDLSNLSIEA 480
 QY 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540
 DB 481 PYGHOASFIISPRACFGWRFAVAEMKAFVTLRRVQEPFIISHPEYEHITLIISRPRIV 540
 QY 541 GREKEGYQMRLOVKPVE 557
 DB 541 GREKEGYQMRLOVKPVE 557

RESULT 3
 US-09-144-367-2
 ; Sequence 2, Application US/09144367
 ; Patent No. 6432639
 ; GENERAL INFORMATION:
 ; APPLICANT: Lichter, Jay
 ; APPLICANT: Guido, Marco
 ; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
 ; FILE REFERENCE: SEQ-12P
 ; CURRENT APPLICATION NUMBER: US/09/144,367
 ; CURRENT FILING DATE: 1998-08-31
 ; PRIOR APPLICATION NUMBER: 60/058,612
 ; PRIOR FILING DATE: 1997-09-10
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 503
 ; TYPE: PRT
 ; ORGANISM: H. sapiens
 US-09-144-367-2

Query Match 13.9%; Score 396.5; DB 4; Length 503;
 Best Local Similarity 28.7%; Pred. No. 5.8e-31;
 Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSN--ASTAFSLYLAPRRS--SLYNLOG--PNHTNY--FTGNFLDILSARTG-----E 61
 DB 7 LANETWLLLAVALSVLLLYLGYTHSHGLPKKIGIPQTPPLPLGN---ILSYHKGFCHMEDNE 63
 QY 62 EHAKYREKYGSTLRAGTAGAPVLNSTDPKVFNHVM--KEAYDY---PKPGMAARVLRAT 117
 DB 64 CHKYGKVMGF---YDG--QQPVLAITDPMIKTVLKECYSVFTNRRRPFPGVGFMKSA-- 117
 QY 118 GGVVTAEGSAHKHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAAEKOMAVGES 177
 DB 118 ---ISIASDEEWKRLRSLSTFTSGKLKENVPITAGYGDVLRNLREA----- 164
 QY 178 AGEKKAATRLT--EGVDVDDKDWVGRATLDVMAAGFDYKSDSLQNKTNELVYVAVGLTDGFA 236
 DB 165 -----ETGKPVTLKDVFGAYSDMVTITSTSGVNIDSLNNPD----- 201
 QY 237 PILDSEFKAIM--WDFV--PYFRTMK--RRHEIPLTQGLAV---SRRVGIELMEQKQAVLGS 290
 DB 202 PFVNTKLLRDFDLDPDFLSITVFPFLLIPLEVLNLCVFPFVNTNLRKSVKR-----M 256
 QY 291 SQAQVDDKDWVGRATLDVMAAGFDYKSDSLQNKTNELVYVAVGLTDGFA 350
 DB 257 KESRIEDTKGHRVDFLQLMIDSONSKETESHKALSDLELVAQSIIFIFAGYEITSSVLSF 316
 QY 351 MFHRLSEDKAVQDKLRBEICQI--DTOMPTLDELNALPYLEAFVKESLRDPPSPYANRE 408
 DB 317 INVETATHPDVQOKLQEEBIDAVLPNKAPPTDYDTLQMEYLDVMVNETLRLLFFIAMRLERV 376
 QY 409 CLKDEDFIPLAEFVIGRDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEFEPRERWL 468

377 CKDVE-----INGMEIPKGVWVMPISYALHRDP-KYWTEPEKFLPERFS 420
US-08-457-274A-25
469 EDVTDLSNS-LEAPYGHQASFGISGRACGWRFAVAEMKAEFLVTLRRVQPEP 520
421 KKKNDIDPIYTPFG-----SGPRNCIGMRPALMMKALIRVLQNFSPK 467
RESULT 4
US-08-457-274A-25
Sequence 25, Application US/08457274A
Patent No. 5734086
GENERAL INFORMATION:
APPLICANT: Scott, Jeffrey G.
APPLICANT: Tomita, Takashi
TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,274A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: Rat
STRAIN:
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
US-08-457-274A-25
Query Match 13.0%; Score 370; DB 1; Length 504;
Best Local Similarity 26.2%; Pred. No. 2.7e-28;
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;
QY 7 LTGALGLAASFWSAIFPSLY-LAPRRSSLYNLOG-----PNHTNYFTGNF-LDI 54
DB 3 LLSALTLETWLLAVLVLLYGFGRTHGLFKQGPCKPLPFFGVLNYYMGLWKFDV 62
QY 55 LSARTGEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFHVM-KEAY-----DYPKP 106
DB 63 -----ECHKKYKIGW---LFDG--QMLPFAITDTEMIKKVLKCFSVFTNRRDFGPV 111
QY 107 GMAARVLRIATGDGVVTAEGSAHKRHRIRIMPSLSAQAVKSMVPIFLEKGMELVDKMD 166
DB 112 GI-----MGKAVSAVAKDEWKRYRALLSPFTTSGLKXEMPIIEQYGDILVKYKQE 163
QY 167 AABKDMAYGESAGEKKAATRLT-EGVDVKWVGRATLDVWALAGFDYKSDSLQN----- 219
Db 164 A-----ETGKEVTMKKVFSGYSMDVITSTFGVNVDSLNNPKDPFV 204
QY 220 -XTNEL-----YVAFVGLTDGFAPTLDSFKAIMW--DFVPYFRMTKRHEIPL 264
Db 205 EKTKKLLRDFDPLFLSVLFPFLT---PIYEMLNICMPFKDSIEFFK----- 250
QY 265 TOGLAVSRVRGIELMEQKQAVLGASDAQVDKDVQGRDILSLVRIANIANLPESO-K 323
Db 251 -----KFVYRMKETRLDSVQKHRV-----DFQLMMNAHNSDKKESHTA 290
QY 324 LSDDEVLAQISNLLFAGYETSSTVLTMFHRLSBKAVQDKLRBEICOI--DTDMPTLDE 381
Db 291 LSDMEITAQSIIFIPAGYEPTSTLSFVLHSLATHPDTQKKLQEIADRALNKAPPTYDT 350
QY 382 LNALPYLEAFVXESLRLDPPSPYANRECKLSDDFIPLAEPIVIGRDGVSINEVRITKGMV 441
Db 351 VMEMEYLDVNLNETLRLYPIGNRLERLVCKKQVE-----INGVFMPKGSVV 395
QY 442 MLPLFNINRSKFTYGEDAEFEFFPERKLEBDVTDSLNS-IEAPYGHQASFGISGRACGWR 500
Db 396 MIPSYALHRDPQHWPE-PEEFPERPSKENKGSIDPYVILPFG-----NGPRNCIGNRF 448
QY 501 AVAEMKAEFLVTLRRVQPEP 520
Db 449 ALMMYKALTKVLQNFSPQ 468
RESULT 5
PCT-US95-05758-25
Sequence 25, Application PC/TUS9505758
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: Cytochrome P450lpr Gene and Its
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05758
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: Rat
STRAIN:
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:

CHROMOSOME/SEGMENT: Unknown
PCT-US95-05758-25

Query Match 13.0%; Score 370; DB 5; Length 504;
Best Local Similarity 26.2%; Pred. No. 2.7e-28;
Matches 147; Conservative 84; Mismatches 189; Indels 140; Gaps 23;

QY 7 LTGALGAFAFASWASIAFFSLY-LAPRRSLYNLQ-----PNTNYFTGTF-LDI 54
DB 3 LLSALTLETWLLAVLVLYGFGTTHGLFKKQIPGPKLPFGFTVLYNGLWKFV 62
QY 55 LSARTGEHAKYREKYGSTLRFAGIAGAPVLNSTDPKVFNMV-KEAY-----DYPKP 106
DB 63 -----ECHKYGKIGW---LEDG--QMPLFAITDTEIMKNVLKCEFSVFTNRDGPV 111
QY 107 GMAARVLRATGCVVTAAGBAKHRRIMISLSAQAVKSWPIFLKGMELVDKXMED 166
DB 112 GI-----MGKAVSVADEERKRYRALLSTFTSGRLKEMFPIIIOYGDILVKYLQ 163
QY 167 AAEKDMAVGESAGEKATRLT-EGVDVKDWVGRATLDVMAAGDPYKSDSLON----- 219
DB 164 A-----ETGKPVTKKVFAGYSMDVITSTSGVNVDSLNNPKDPFV 204
QY 220 -KNEL-----YVAVGLTGDGAPTLDSFKALMW---DFVPYFTMKRRHSLPL 264
DB 205 ETKKLLRDFDPLFLSVLPFFLT-----PIYEMLNICMFPKDSIEFFK----- 250
QY 265 TQGLAVSRVGIEMLEKQKQAVLGSASDAQVDKQVQGRDILSLVRANIANLPESQ-K 323
DB 251 -----KPVYRMKTRLDVSQHRV-----DFQLMNAHNSDKESHSTA 290
QY 324 LSDEEVLAQISNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREEICQI--DTDMPTLDE 381
DB 291 LSDMEITAOQIIFAGYBPTSTLSVLHSLATHFDTKQKLQEEIDRALPNKAPPTYDT 350
QY 382 LNALPYLEAFVKESLRLDPPSPYANRECKDDEFTPLAEPVIGRDSGVINEVRITKGMV 441
DB 351 VMEVEYLDWVNTLRLYFIGNLERVCKDVE-----INGVFWPKGSVV 395
QY 442 MLPLFNINRSKFIYGBDAEERPERWLEDVTDLSNS-TEAPVGHQASISGPRACFGW 500
DB 396 MIPSYALHRDPQHWPE-PSEFRERPSKENKSIDPVVYLPFG-----NGFRNCIGMR 448
QY 501 AVAEMKAFVTLRRVQFEP 520
DB 449 ALMMKALTKVLQNFSPQ 468

RESULT 6
US-09-527-073-2
; Sequence 2, Application US/09527073
; Patent No. 6534313
; GENERAL INFORMATION:
; APPLICANT: Michael M. Neff
; APPLICANT: Joanne Chory
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
; FILE REFERENCE: SALKINS.024
; CURRENT APPLICATION NUMBER: US/09/527,073
; CURRENT FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170,931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172,832
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

US-09-527-073-2

Query Match 12.9%; Score 367; DB 4; Length 520;
Best Local Similarity 22.2%; Pred. No. 5.7e-28;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

QY 4 LVLITGALGAFAFASWASIAFFSLYLAARR-----SSLYNLQGNHNTYFTGTFDILS--- 56
DB 13 VLVLSVLVLVVKGMSL-----LWRRPKIEEHFSKQIGRPPY-HFFIGNVKELVGMML 67
QY 57 -----ARTCEEHAKYREKYGSTLRFAGIAGAPVLNST--DPKVFNMVHVKEA 100
DB 68 KASSHPMPFSNINLPRVLSFYHWRKIYGATFL---VWFGTFTLTVDADDLIREFSKS 124
QY 101 YDYPFGMAARVLRATGCVVTAAGBAKHRRIMIPSLSAQAQVSMWPIFLKGMELV 160
DB 125 BFEYK-NEAHLPLVKQLEGDLGLSLKGEKWAHHRKIIISPTFHMENKLLVPLVWLSVTD 183
QY 161 DQWEDAAEKDMAVGESAGEKATRLTETGVVDKDWVGRATLDVMAAGDPYKSDSLQNK 220
DB 184 DKMSDKLSN-----GEVE-----VDVYFENFQILTEDVISTAFGSSYE----- 222
QY 221 TNELVAVFVGLTDGFAP-----TLDSFKALIMWDFVPYFTMKRRHEIFLTQGLAV 270
DB 223 -----DGEAVFRLQAQQMLLCAEAFQV---FIPGYRFPF-----TRGNLX 260
QY 271 SRRVG-----IEMEQKQAVLGSASDAQVDKQVQGRDILSLVRANIANLPESQK 323
DB 261 SKLDEKEIRKSLKLIERRQNAIDGEBEC---KEPAKDLGLGLMIOA-----KN 308
QY 324 LSDEEVLAQISNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREEICQI--DTDMPTLDE 381
DB 309 VTVDIVIECKSFYFAGKQTTSNLLTWTLLSMHPEWQAKARDEVLRVCGSRDVPDKH 368
QY 382 LNALPYLEAFVKESLRLDPPSPYANRECKDDEFTPLAEPVIGRDSGVINEVRITKGMV 441
DB 369 VVYKLTLSMILNESURLYPPVIVATIRRAKSDVK-----LGGYKIPQGTTEL 413
QY 442 MLPLFNINRSKFIYGBDAEERPERWLEDVTDLSNS-TEAPVGHQASIS---GPRACFGW 498
DB 414 LPIITAVHDDQAINWDVNEFNPAFADGVPRAAK-----HPVGFIPGLGVRTTCIG 466
QY 499 RFPAEMKAFVTLRRVQFEPFIISHPEYEHITLII 534
DB 467 NLAILQAKLTAVMIQRTFTH---LAPTYQHAPTVL 499

RESULT 7
US-09-976-594-533
; Sequence 533, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 533
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 6825202CD1
; US-09-976-594-533

Query Match 12.1%; Score 345; DB 4; Length 524;
Best Local Similarity 23.4%; Pred. No. 9.4e-26;
Matches 140; Conservative 104; Mismatches 207; Indels 148; Gaps 25;


```
QY 172 MAVGSAGEKATRLTEGVYKDVWGRATLDVMAAGFDYKSDLSQNKTNELVAVGL 231
Db 192 F-----KHVRKRGQTFDIOELFFRLTVDSATEFLFGSASLRDSD-----VGL 236
QY 232 TDGFAPTLDSFKAIMWDFVYFERTK--RRHEIPITQGLAVSRRYGIELMPEQKQAVLGS 289
Db 237 T-----PTTKDFEG-RGDFADAFNSQYQAYRFLQKQWILN--GAFF--RKSIIVHK 287
QY 290 ASQAVDK-----KDVQRDILSLVRANIANLPESQKLSDEVLA-QISNLLFAGYE 342
Db 288 FADHYVQKALELDDDLQKQDGVFELYEL-----AKQTRDPKVLDRDQNLNLVAGRD 339
QY 343 TSSTVLTVMFHLRSEDKAVQDKLREI-----CQIDTDMPTLDLALNLPYLAFAFKES 395
Db 340 TTAGLSFVYFELSRNPEVFAKLREENVNRFGLGEEARVEEISFESLAKSCEYLKAVINEA 399
QY 396 LRLDPPSPVANRECLKDEDFIPLAEVIGRDSVINEVRITKGTVMPLPLFNINRSFYI 455
Db 400 LRLYPSVPHNFRVATRNTT-LPRGG---GKDG--CSPIVVKKGQVVMYTV-GTHRDPSTY 453
QY 456 GEDAEFRPERWLEDVTDLSNLSIAPYGHQASFGSPRACFGWRFAVAKMKAFLFVTLRR 515
Db 454 GADADVFRPERWEPETRKLWAYVP-----FNGGPRICLQGPALTEAS---YTVRL 504
QY 516 VQ-----PEPIISHPEYEHITLIIS 535
Db 505 LQEFGNLSLDPNAEYPPKLQNTLTL 530
```

RESULT 10

```
US-08-457-274A-22
; Sequence 22, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Musca domestica
; STRAIN: Rutgers
```

```
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: Chromosome 5
; US-08-457-274A-22
Query Match 11.6%; Score 330.5; DB 1; Length 507;
Best Local Similarity 23.5%; Pred. No. 2.6e-24;
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;
QY 10 ALGLAFAWASIAFSLYLAARRSSLYNLQGNH--TNYFTGFLDLSA-RIGEBHAKY 66
Db 10 ALGV-----LASLALYFVRWFGYWKRRGIPHEEPHLVMGNVKGSRKSHIGELIADY 62
QY 67 RKYGSTLRFAGI--AGAPVLNSTDPKVFNVHM--KEAYDYPKPGMAARVLRIATGQVVT 123
Db 63 YRKFKSGPFGAGIFLGHKPAAVVLDKELRKRVLLKDPNFANRGLYNYNEKDDPLTGHVLM 122
QY 124 ARGEAHKRRHRIIMPSSLAQAVKSNVIFLEKGMELVDKMDAAEKDMVAGSAGEKKA 183
Db 123 VEGERWSLRTKLSFTFTAGKMYNTVLEVGORLLEVME----- 164
QY 184 TELETGEG-VDVKDWYGRATLDVMAAGFDYKSDLSQNKTNELYV 226
Db 165 -KLEVSSSLDMDILARENTDVGSAFGIECSLRNPHDRFLAMGRKSTEVPRHNALIM 223
QY 227 AFVGLTDGFAPTLDSFKAIMWDFVYFERTKRRHEIPLTQGLAVSRRYGIELMPEQKQAV 286
Db 224 AFI-----DSFPEL-----SRKLGMRVLPEDEVHOF 248
QY 287 LGSASDAQVD---KKDVQGRDILSLVRANIANLPESQK-----LSDEEVLQGISNLLFA 339
Db 249 FWSISKEVVDYREKKNIRNDFLDLVDLK---NNPESISKGLGLTFNELAAQVVFVFLG 305
QY 340 GYETSTVLTVMFHLRSEDKAVQDKLREI-----CQIDTDMPTLDLALNLPYLAFAFKES 395
Db 306 GFETSSSTMTGFALYELAQNLQDLRLREENVNEDQFKEDNISYDALMNPYLDQVLNET 365
QY 396 LRLDPPSPVANRECLKDEDFIPLAEVIGRDSVINEVRITKGTVMPLPLFNINRSFYI 455
Db 366 LRKYVGSALTQTLNDY-----VYPHNPKYV---LPKGTLVIFVPLGIHYDPELY 413
QY 456 GEDAEFRPERWLEDVTDLSNLSIAPYGHQASFGSPRACFGWRFAVAKMKAFLFVTLRR 515
Db 414 -PNPEFDPERESPVMKQSDVD-----WLFGFGDGPNCIGMRFKQVQSRGLGLALVIRH 467
QY 516 VOFE-----PLISHPE 526
Db 468 FRFTVCSRTDIPMQINPE 485
```

RESULT 11

```
PCT-US95-05758-22
; Sequence 22, Application PC/TUS9505758
; GENERAL INFORMATION:
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05758
; FILING DATE:
```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
STRAIN: Rutgers
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 5
PCT-US95-05758-22

Query Match 11.6%; Score 330.5; DB 5; Length 507;
Best Local Similarity 23.5%; Pred. No. 2.6e-24;
Matches 131; Conservative 94; Mismatches 210; Indels 123; Gaps 19;

QY 10 ALGUAFAFWASIAFFSLYLAPRSSLNLOQPH--TNFTGNFLDILSA-RTGEHAKY 66
DB 10 ALGV-----LASLALFYFWMFNGYKRRGIPHEEPHLMVGNVKKSLRKYHIGIADY 62
QY 67 REKVGSTLRPAGI--AGAPVLNSTDPKVENHVM--KEAYDPKPGMAARVLRIATGDGVVT 123
DB 63 YRKFGSGPAGIFLGHKPAAVLDKELRXLVIKDFSNFANRGLYNEKDDPLATGLVM 122
QY 124 AGEPAHRHRIMTIPSLSAQAVKSMVPIFLEKGMELVDKMDAIAEKMDAVGESAGEKAT 184
DB 123 VEGEKWRSRLTKLSFTTAGMKMTYNTVLEVGQRLLEVNYE-----KTNELV 226
QY 184 TRLETEG--VDYKDVGRATLDVMAAGFDYKSDSLQN-----KTNELV 226
DB 165 -KLEVSSELDWRDILARNFDVIGSAVAGIECSLNRPHDRFLAMGKKSIEVPRNALIM 223
QY 227 AFVGLTDFAPTLDSFKAIMWDFVYFRTMKRRHEIPLTQGLAVSRVRIELMEQKQAV 286
DB 224 AFI-----DSFPEL-----SRKLGRVLPEDVHQF 248
QY 287 LGSASDAQVD---KDVQGRDILSLVRYANIAANLPESQK-----LSDEEVLQISNLLEFA 339
DB 249 FMSSTKETVDYREKNNIRNDFDLVLDLK---NNPESISKGLGLTFNEAAQVTFVFLG 305
QY 340 GYETSSTVLTWFRHLSKAVQKLEEEI---CQDITDMPITLDELNALPYLEAFVKES 395
DB 306 GFETSSTMGALYELAQNLQQLDRLEEVNEVDFQKEDNISYDALMNIPLYDQVLNET 365
QY 386 IRLDPPSYANRECKDEDFIPLAEVPIGRDGSVINEVRITKGTMMWMLPLFNINRSKFIY 455
DB 366 LRKYPVGSALTROTINDY-----VVPHPKV-----LPKGTLPFIPVLGHIHPDELY 413
QY 456 GDAEERPERLEWDTVDSLNSIEAPYGHQASFTSGPRACFGWRFAVAEMKAFPLVTLRR 515
DB 414 -NPPEDFPEPSMVKQRSVD-----WLFGDGRPCNGICGRFQWQSRGLGALVIRH 467
QY 516 VQFE-----PIISHPE 526
DB 468 PRFTVCSRTDIPMQINPE 485

Query Match 11.5%; Score 328; DB 4; Length 504;
Best Local Similarity 27.1%; Pred. No. 4.6e-24;
Matches 127; Conservative 75; Mismatches 185; Indels 82; Gaps 18;

QY 68 EKYGSTLRFAGI--AGAPVLNSTDPKVENHVM--KEAYDPKPGMAARVLRIATGDGVVTA 124
DB 66 ERYG---EMWGLYEGQQPMLVMDPMIKTVLKECYSVFTNQPLGPMGLK-SALSFA 121
QY 125 EGEAHRHRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMDAIAEKMDAVGESAGEKAT 184
DB 122 EDEEMKRIITLSPAPTSVKFKEMVPIISQCCDMLVRSLOEAE----- 165
QY 185 RLETEGVVDKDVGRATLDVMAAGFDYKSDSLQNTNELYVAFVGLTDGFPATLDSFKA 244
DB 166 --NSKINLKDFFGAVTMDVITGTLFGVNLDSLNNPQ-----PFLKWKX 209
QY 245 IM-WDEVPVFTMKRRHEIPLTQGLA--VSRVGIELMBQKQAVLGSASDAQVDK--KDV 300
DB 210 LKLLDPLDFELL-----ISLFFPLTPVFEALNIGLFPKDVTHFLKNSIERMKESRLKDK 264
QY 301 QGR--DILSLVRYANIAANLPESQKLSDEEVLQISNLLEFAGYETSSTVLTWFRHLS 358
DB 265 QHRVDFVFOQMDSQNSKETKSHKALSDLELVAQSIIFAIAYDTTSTLPIFYELATH 324
QY 359 KAVQDKLRREICQIDTDMF-----TIDELNALPYLEAFVKESRLDPPSYANRECKDE 413
DB 325 PDVQKQLEH---IDAVLPNKAPVTVDALVQNEYLDMVNETLRLFPVVSRTVCKKDI 381
QY 414 DTIPLAEVPIGRDGSVINEVRITKGTMMWMLPLFNINRSKFIYGEDAEBFRPE-RWLEDVT 472
DB 382 E-----INGVFIPKGLAVMVPITYALHHDH-KYWTPEFKPCPSERFSKKNK 425
QY 473 DSLNSIE-APYGHQASFTSGPRACFGWRFAVAEMKAFPLVTLRRVQFEP 520
DB 426 DSIDLRYIYFPG-----AGPRNCIGWRFALTNIKLAIVIRALQNFSEFKP 468

RESULT 13
US-09-302-620B-98
Sequence 98, Application US/09302620B
Patent No. 6334420
GENERAL INFORMATION:
APPLICANT: Wilson, C. Ron
APPLICANT: Craft, David L.
APPLICANT: Birch, Dudley
APPLICANT: Behoo, Mark
APPLICANT: Madduri, Krishna M.
APPLICANT: Cornett, Cathy A.
APPLICANT: Brenner, Alfred A.
APPLICANT: Tang, Maria
APPLICANT: Loper, John C.
APPLICANT: Gleeson, Martin
TITLE OF INVENTION: CYTOCHROME P450 MONOOXYGENASE AND NADPH CYTOCHROME P450
OXIDOREDUCTASE GENES AND PROTEINS RELATED TO THE OMEGA
TITLE OF INVENTION: HYDROXYLASE COMPLEX OF CANDIDA TROPICALIS AND METHODS

; TITLE OF INVENTION: RELATING THERETO
 ; FILE REFERENCE: 1010-16 seq
 ; CURRENT APPLICATION NUMBER: US/09/302,620B
 ; CURRENT FILING DATE: 1999-04-30
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 98
 ; LENGTH: 540
 ; TYPE: PRT
 ; ORGANISM: Candida tropicalis
 ; US-09-302-620B-98

Query Match 11.5%; Score 328; DB 4; Length 540;
 Best Local Similarity 23.9%; Pred. No. 5.1e-24;
 Matches 133; Conservative 93; Mismatches 197; Indels 134; Gaps 22;
 QY 42 NHTNYP-----TCNFDILSARTGEB--HAK-----YREKYG 71
 DB 25 NYTRWYFPLVLLSUNFISLHTRYLERRRFAKPLGNFVDPDPTFGIATPLLLIYLSKG 84
 QY 72 STURFA-----GIAGAPVLNSTDPKVENHVMKEAYDYPKPGMAAR 111
 DB 85 TVNKPFWGLWNNKYIVRDPKYKTTGLRIVGLPLIETMPENIKAVLATQFNDFSLGTRHD 144
 QY 112 VLIATGDDVVTAGBAHGRHRIMIPSLSAQAVKSMVPIFEKGMELVDQMDEAAEKD 171
 DB 145 FLYSLLDGDIPTLDGAGWKHSRTMLRPOPAREQVSHV-----KLEBPHVQVF 191
 QY 172 MAVGESAGKATRLTEGVGVKMDVGRATLDVMALAGFDYKSDSLQNKTNELVAVFVGL 231
 DB 192 F-----XVYKRGQTDFIQELFFRLTVDSATEFLFGESAESLRDES-----IGL 236
 QY 232 TDGFATLDSFKAMWDFVYFRTWK--RRHEIPTQGLAVSRVRVIGIELMEQKQAVLGS 289
 DB 237 T-----PTTKDFDG--RRDFADAENYSQYQAVRFLQOMYILN--GSEF--RKSIAVYHK 287
 QY 290 ASDQAVDK-----KDVQGRDILSLVVRANIANLPESOKLSDEEVL--QISNLLFAGYE 342
 DB 288 FADHYVQKALELTDQLQDQGVFLYEI-----AKQTRDPKVRDQLNLVLVGRD 339
 QY 343 TSSTVLTWFMHRLSEDAVQDKLREPI-----CQIDTDMPTLDELNALPYLEAFVKES 395
 DB 340 TTAGLLSFVYELSRNPEVFAKLREVENRFGLGSEARVEEISFESLAKSCVYLKAVINET 399
 QY 396 LRDPSPYANRECLDKDEFIPLAEVPIGDSVINEVRITKGTWMLPLFNINRSKFIY 455
 DB 400 LRLYPSVPHNFRVATNTT-LPRGG--GEDG--YSPIVVKGVQVMVTVIATHRPSIY 453
 QY 456 GEDAEPRPERWLEDVTDLSNLSIEAPYGHQAFISGPACFGWRFAVAKMKAFLFVTLRR 515
 DB 454 GADADVPRPERWEPETRKLWAYVP-----FNGGPRICLGGQFALTEAS---YVTVRL 504
 QY 516 VQEPFIISHPEYDHITL 532
 DB 505 LQ-----EFAHLSM 513

RESULT 14
 US-09-158-767-19
 ; Sequence 19, Application US/09158767A
 ; Patent No. 6180363
 ; GENERAL INFORMATION:
 ; APPLICANT: Batard, Yannick
 ; APPLICANT: Durst, Francis
 ; APPLICANT: Schalk, Michel
 ; APPLICANT: Werck-Reichhart, Daniele
 ; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
 ; FILE REFERENCE: A32000
 ; CURRENT APPLICATION NUMBER: US/09/158,767A
 ; CURRENT FILING DATE: 1998-09-23
 ; EARLIER APPLICATION NUMBER: FR 97-12094
 ; EARLIER FILING DATE: 1997-09-24

; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 19
 ; LENGTH: 541
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Altered sequences
 ; US-09-158-767-19

Query Match 11.4%; Score 326.5; DB 3; Length 541;
 Best Local Similarity 24.6%; Pred. No. 7.3e-24;
 Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;
 QY 78 GIA--GAPVLNSTDPKVENHVMKEAYD--YKPGMAARVLRIATGDDVVTAGBAHGRHR 134
 DB 73 GVARRGGVLVTCTDPNLEHLVKARFDNYPKPGFWHGVFEDLLGDI FNSDGDGTWLAQRK 132
 QY 135 IMIPLSAQAVKSMVPIFEKGMELVDKMDAAEKDMAVGESAGKATRLTEGVGVK 194
 DB 133 TAALEFTRTTRTAMSRWVSRSH--GRLLPILA--DAAKGKQAQ-----VDLQ 176
 QY 195 DWVGRATLDVMALAGFDYKSDSLQ--KTNELVAVFVGLTDGFPATLDSFKAIMWDFVPY 252
 DB 177 DILLRLTFDNICGLAFGKDPETLAQGLPENEFASAFDTEA--TLNRF-----IFPE 227
 QY 253 F-----RTMKRRHEIPTQGLA--VSRVIGIELMEQKQAVLGSASDAVDKDVQGRDIL 306
 DB 228 FLWRCKKWLGLMETTTSSMAHYDQYLAIVIKRLELAAGNGKCDTAAATHD-----DLL 283
 QY 307 SLLVRANIANLPESOKLSDEEVLQISNLLFAGYETSTVLTWFMHRLSEDAVQDKLR 366
 DB 284 SPMFKG-----SYSDSLQHVNLFLAGRDTSSVALSWFFVLVSTHPAVERKIV 334
 QY 367 EIIQI-----DIDMETLDELNALPYLEAFVKESLRLDPPSPYANRECLKDE 413
 DB 335 RELCSVLAASRGADHPALMLAEPTFBEOLDRLVYLKALSETLRLYPSVPEDSKHVVAD- 393
 QY 414 DIPLAEVPIGDSGVINEVRITKGTWMLPLFNINRSKFIYGEDAEFEPRERWLE---- 469
 DB 394 DYLP-----DGTFF-----VPAGSSVYIYISAGRMKGWGEDCIEFPERWLSADGT 440
 QY 470 --DVTDSNLSIEAPYGHQAFISGPACFGWRFAVAKMKAFLFVTLRRVQEPFIISHPEY 527
 DB 441 KPEQDSYKVF-----AFNAGPRVCLGKDLAYLQMKNIAGSVLLRHLRTVAPGRHVE 492
 QY 528 EHITLISRPRVIGREKEGYQVRLQVKE 555
 DB 493 QKMSLTFLM-----KGG--LRMEVRP 511

RESULT 15
 US-09-158-767-20
 ; Sequence 20, Application US/09158767A
 ; Patent No. 6180363
 ; GENERAL INFORMATION:
 ; APPLICANT: Batard, Yannick
 ; APPLICANT: Durst, Francis
 ; APPLICANT: Schalk, Michel
 ; APPLICANT: Werck-Reichhart, Daniele
 ; TITLE OF INVENTION: RECODING OF DNA SEQUENCES PERMITTING
 ; FILE REFERENCE: A32000
 ; CURRENT APPLICATION NUMBER: US/09/158,767A
 ; CURRENT FILING DATE: 1998-09-23
 ; EARLIER APPLICATION NUMBER: FR 97-12094
 ; EARLIER FILING DATE: 1997-09-24
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 541
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: Altered sequences
US-09-158-767-20

Query Match      11.4%; Score 326.5; DB 3; Length 541;
Best Local Similarity 24.6%; Pred. No. 7.3e-24;
Matches 125; Conservative 87; Mismatches 197; Indels 99; Gaps 20;

Qy 78 GIA--GAPVLNSTDPKVFHVMKEAYD-YPKPGMAARVLRIATGDGVVTABGEAKHRHR 134
Db 73 GVARRGLVTVTCDPRNLEHVKARFDNYKGFVHGVFDLLGDGIFNSDGDWTWLAQRK 132
Qy 135 IMPPSLSAQAVKMWPIFLKGMELVDKMDAEDKMDMAGVSGAGEKAKATRLTEGVYK 194
Db 133 TAALEFTRTLRTAMSRWVSRSH--GRLLPILA--DAAKGKAQ-----VDLQ 176
Qy 195 DWVGRATLDVMALAGFDYKSDSLQN--KTNELYVAFVGLTDGFAPTLDSFKAIMWDFVY 252
Db 177 DLLRLTFDNLICGLAFGKDEPTELAQGLPENEFASAFDRATEA---TLNRF-----IFPE 227
Qy 253 F-----RTMKRRHEIPLTQGLA--VSRRVGIELMEQKQAVLGSASDAQVDKDKVOGRDIL 306
Db 228 FLWRCKKWLGLGMETTTLTSSMAHVDQYLAAVIKRKLKLELAAGNGKCDTAATHD----DLL 283
Qy 307 SLLVRANIAANLPESQKLSDEEVLQISNLPAGYETSTVLTWMPHRLSEDKAVQDKLR 366
Db 284 SRFWRKG-----SYDESIOHVALNFILAGRTSSVALSWFPLVSTHPAVERKIV 334
Qy 367 EEICQI-----DTMPTLDELNALPYLEAFVKESLRLDPPSPYANRECLKDE 413
Db 335 RELCSVLAASRGADHPALWLAESPFTFEELDRVLVTKAALSETLRLYPSVPEDSKHVAD- 393
Qy 414 DFIPLAEPVIGRDSVINEVRIITKGTWMLPLFNINRSKIYGEDAEERPERWLE---- 469
Db 394 DYLP-----DGTG-----VPAGSSVTYSIISAGRMKGWGEDCLEFRPERWLSADGT 440
Qy 470 --DVTDSLNSIEAPYGHQASFGISGFACFGWRFAVAKMKAFLFVTLRRVQFEPFISHPEY 527
Db 441 KFEQHDSYKFV-----AFNAGPRVCLGKLAYLQMKNIAGSVLLRHRLTVAPGHRVE 492
Qy 528 EHITLISRPRIVGREKEGYQMRLOQXP 555
Db 493 QKMSLTLM-----XGG--LRMEVRP 511
```

Search completed: April 2, 2004, 14:02:16
Job time : 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 14:00:49 ; Search time 40 Seconds
(without alignments)
3653.561 Million cell updates/sec

Title: US-10-066-007-3
Perfect score: 2852
Sequence: 1 MFILVLTGALGAAPSWAS.....RIVGREKEGYQRLQKPVVE 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1070241 seqs, 262374223 residues

Total number of hits satisfying chosen parameters: 1070241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	100.0	557	14	US-10-066-007-1
2	2852	100.0	557	14	US-10-066-007-3
3	398.5	14.0	503	15	US-10-313-963A-56
4	396.5	13.9	503	14	US-10-146-575-2
5	385.5	13.5	503	9	US-09-957-997-3
6	384	13.5	537	12	US-10-425-114-38180
7	377.5	13.2	527	12	US-10-425-114-67055
8	376.5	13.2	547	12	US-10-425-114-65616
9	368.5	12.9	547	12	US-10-425-114-61218
10	368.5	12.9	560	12	US-10-425-114-59349
11	368.5	12.9	562	12	US-10-425-114-59350
12	367	12.9	520	9	US-09-992-901-2
13	365.5	12.8	622	12	US-10-425-114-38827
14	365	12.8	520	12	US-10-114-270-60
15	333	12.4	524	12	US-10-424-599-199559

16	352	12.3	533	12	US-10-424-599-269344	Sequence 269344, App
17	350	12.3	520	12	US-10-072-012-548	Sequence 548, App
18	350	12.3	520	14	US-10-032-189-115	Sequence 115, App
19	349.5	12.3	520	12	US-10-424-599-228819	Sequence 228819, App
20	349	12.2	508	14	US-10-274-694-16	Sequence 16, Appl
21	348.5	12.2	546	15	US-10-291-365-405	Sequence 405, App
22	347.5	12.2	524	15	US-10-291-365-877	Sequence 877, App
23	347.5	12.2	524	15	US-10-291-365-878	Sequence 878, App
24	347.5	12.2	524	15	US-10-291-365-879	Sequence 879, App
25	346.5	12.1	524	9	US-09-989-722-264	Sequence 264, App
26	346.5	12.1	524	9	US-09-989-723-264	Sequence 264, App
27	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
28	346.5	12.1	524	9	US-09-989-727-264	Sequence 264, App
29	346.5	12.1	524	9	US-09-989-731-264	Sequence 264, App
30	346.5	12.1	524	9	US-09-989-732-264	Sequence 264, App
31	346.5	12.1	524	9	US-09-991-073-264	Sequence 264, App
32	346.5	12.1	524	9	US-09-990-442-264	Sequence 264, App
33	346.5	12.1	524	9	US-09-991-163-264	Sequence 264, App
34	346.5	12.1	524	9	US-09-993-604-264	Sequence 264, App
35	346.5	12.1	524	9	US-09-990-456-264	Sequence 264, App
36	346.5	12.1	524	9	US-09-989-721-264	Sequence 264, App
37	346.5	12.1	524	9	US-09-992-598-264	Sequence 264, App
38	346.5	12.1	524	9	US-09-989-293A-264	Sequence 264, App
39	346.5	12.1	524	9	US-09-989-735-264	Sequence 264, App
40	346.5	12.1	524	9	US-09-990-444-264	Sequence 264, App
41	346.5	12.1	524	9	US-09-991-181-264	Sequence 264, App
42	346.5	12.1	524	9	US-09-989-730-264	Sequence 264, App
43	346.5	12.1	524	9	US-09-990-438-264	Sequence 264, App
44	346.5	12.1	524	9	US-09-993-687-264	Sequence 264, App
45	346.5	12.1	524	10	US-09-989-734-264	Sequence 264, App

ALIGNMENTS

RESULT 1

US-10-066-007-1
; Sequence 1, Application US/10066007
; Publication No. US2003007691A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/10/066,007
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US/09/518,386
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: TRANSIT
; LOCATION: (1)...(26)
US-10-066-007-1

Query Match 100.0%; Score 2852; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.3e-248;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFILVLTGALGAAPSWASIAFFSVLAPRRSLNLOQPNHNTYFTGNFDILSARTG 60

Db 1 MFILVLTGALGAAPSWASIAFFSVLAPRRSLNLOQPNHNTYFTGNFDILSARTG 60

QY 61 BEHAKYREKYGSLRFRAGIAGAPVLNSDTPKVFNVHNVKGAIDYKPCGMARVLRIATG 120


```

Db 61 EEAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNVHKEAYDYFKPGVAARVLRATGDG 120
QY 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAGSAGE 180
Db 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAGSAGE 180
QY 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELNLYVAFVGLTDCGFAPTLD 240
Db 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELNLYVAFVGLTDCGFAPTLD 240
QY 241 SFKAIMWDFVYFRTMKRRHEIPITQGLAVSRVGIEMLEQKQAVLGSASDAQVKKDV 300
Db 241 SFKAIMWDFVYFRTMKRRHEIPITQGLAVSRVGIEMLEQKQAVLGSASDAQVKKDV 300
QY 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
Db 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
QY 361 VQDKLREEICQIDTDMPTLDLINALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLINALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
QY 421 PVIGRDSGVINEVRIITKGTWMLPLFNINRSKFYIGEDAEPRPRERWLEDVTDLSINSTE 480
Db 421 PVIGRDSGVINEVRIITKGTWMLPLFNINRSKFYIGEDAEPRPRERWLEDVTDLSINSTE 480
QY 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
Db 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
QY 541 GREKEGYQMRLOVKPVE 557
Db 541 GREKEGYQMRLOVKPVE 557

```

RESULT 2

```

US-10-066-007-3
; Sequence 3, Application US/10066007
; Publication No. US20030077691A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/10/066,007
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US/09/518,386
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 3
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-10-066-007-3

```

```

Query Match 100.0%; Score 2852; DB 14; Length 557;
Best Local Similarity 100.0%; Pred. No. 2.3e-248;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MFILVLITGALGAALFASWASIAFFSLYLAPRRSSLYNLOQPNHTNYFTGNFLDILSARTG 60
Db 1 MFILVLITGALGAALFASWASIAFFSLYLAPRRSSLYNLOQPNHTNYFTGNFLDILSARTG 60
QY 61 BEHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNVHKEAYDYFKPGMAARVLRATGDG 120
Db 61 BEHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNVHKEAYDYFKPGMAARVLRATGDG 120

```

```

QY 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAGSAGE 180
Db 121 VVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAGSAGE 180
QY 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELNLYVAFVGLTDCGFAPTLD 240
Db 181 KKATRLTEGVDDKQWVGSRATLDVMAALAGFDYKSDSLQNKTNELNLYVAFVGLTDCGFAPTLD 240
QY 241 SFKAIMWDFVYFRTMKRRHEIPITQGLAVSRVGIEMLEQKQAVLGSASDAQVKKDV 300
Db 241 SFKAIMWDFVYFRTMKRRHEIPITQGLAVSRVGIEMLEQKQAVLGSASDAQVKKDV 300
QY 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
Db 301 QGRDILSLVRANIANLPSQKLSDEEVLAQISNLLFAGYETSTVLTWTFHRLSDEKA 360
QY 361 VQDKLREEICQIDTDMPTLDLINALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
Db 361 VQDKLREEICQIDTDMPTLDLINALPYLEAFVKESLRDLDPSPYANRECLKDEDFIPLAE 420
QY 421 PVIGRDSGVINEVRIITKGTWMLPLFNINRSKFYIGEDAEPRPRERWLEDVTDLSINSTE 480
Db 421 PVIGRDSGVINEVRIITKGTWMLPLFNINRSKFYIGEDAEPRPRERWLEDVTDLSINSTE 480
QY 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
Db 481 PYGHQASFTSGPRACFGWRFAVAKMAFLFVTLRRVQFEPPIISHPEYEHITLIISRPRIV 540
QY 541 GREKEGYQMRLOVKPVE 557
Db 541 GREKEGYQMRLOVKPVE 557

```

RESULT 3

```

US-10-313-963A-56
; Sequence 56, Application US/10313963A
; Publication No. US20040002078A1
; GENERAL INFORMATION:
; APPLICANT: Boutell, Jonathan
; APPLICANT: Godber, Benjamin
; APPLICANT: Hart, Darren
; APPLICANT: Blackburn, Jonathan
; TITLE OF INVENTION: Arrays
; FILE REFERENCE: KIL-001
; CURRENT APPLICATION NUMBER: US/10/313,963A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/335,806
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/410,815
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 56
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-313-963A-56

```

```

Query Match 14.0%; Score 398.5; DB 15; Length 503;
Best Local Similarity 28.7%; Pred. No. 6.2e-27;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

```

```

QY 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLOQ--PNHTNY--FTGNFLDILSARTG-----E 61
Db 7 LMETWLLLASLVLLYLYGTHSHGLPKLGIPTPLFLGN---ILSYHKGCFMEDME 63
QY 62 BEHAKYREKYSTLRFPAGIAGAPVLNSTDPKVFNVHVM--KEAYDY---PKPGMAARVLRAT 117
Db 64 CHKKYKVMGF---YDG--QQPVLAITDPDMIKTVLKECYSVFTNRRPFGPVGFMKSA-- 117
QY 118 GDCVTAEGEAHKRRIRIMIPSLSAQAVKSMVPIFLEKGMELVDKMMEDAAEKDMVAGS 177

```

```

Db 118 ---ISIAEDDEWKLRLSPTFTSCKLKEMVPIIAQGDVLRNLRREA----- 164
Qy 178 AGEKATRLLET-EGVDVQKWDVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTGDGA 236
Db 165 -----ETGKPVTLKDVFGAYSDMTSTSTSGVNIIDSLNNPQD----- 201
Qy 237 PTLDSFKAIM-WDFV-PYFRMTK-RRHEIPTQGLAV---SRRVGIEMEQKQAVLGSA 290
Db 202 PFVENTKLLRFDFLDPFFLSITVFFFLPILEVLNICVFPREVTFNLRKSVKR-----M 256
Qy 291 SDQAVDKVQVQGRDILSLVLRANIANLPESQKLSDEEVLAQISNLLFAGYETSTVLTW 350
Db 257 KESRLEDTQKRVDFLQMDISQSKETESHKALSDELVAQSIIFIFAGYETTSVLSF 316
Qy 351 MEHRLSEDAVQKLRKEICQI--DTDMPTLDLNLALPYLEAFVKESLRLDPPSPYANRE 408
Db 317 IMVELATHPDVQKQLEIDAVLPNKAPPTDYVLQMEVLDVWVNETLELFFIAMRLERV 376
Qy 409 CLKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPERWL 468
Db 377 CKKQVE-----INGMFIPKGVVMIPSYALHRDP-KYWTEPEKFLPERFS 420
Qy 469 EDVTDLSNS-IEAPYGHQASFIQGRACFGHFAVAEMKAEFLVTLRRVQREP 520
Db 421 KKNKDNIDPIYITPFG-----SGPRNCIGMRFPALMMKMLALIRVLQNFSPRP 467

RESULT 4
US-10-146-575-2
; Sequence 2, Application US/10146575
; Publication No. US20030059800A1
; GENERAL INFORMATION:
; APPLICANT: Guido, Marco
; APPLICANT: Liciter, Jay
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-122
; CURRENT APPLICATION NUMBER: US/10/146,575
; PRIOR FILING DATE: 2002-05-14
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 503
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-146-575-2

Query Match
Best Local Similarity 28.7%; Score 396.5; DB 14; Length 503;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

Qy 13 LAAPSW--ASTAFPSLYLAPRRS--SLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLAVSLVLLYLYGTHSHGLFKLGIPOGTPLEFLGN---ILSYHKGFCMFDM 63
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLSNSTDPKVFNHVM-KEAYD---PKPGMAARVLRIAT 117
Db 64 CHKKYKGVNGF---YDG--QQPVLAITDPDMIKTVLKECYSVFTNRRPFGPVGFMKSA- 117
Qy 118 GDGVVTAEGEAHRRHRIIMPSLSAQAVKSMVPIFLKGMELVDKQMEDAAEKDQAVGES 177
Db 118 ---ISIAEDDEWKLRLSPTFTSCKLKEMVPIIAQGDVLRNLRREA----- 164
Qy 178 AGEKATRLLET-EGVDVQKWDVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLTGDGA 236
Db 165 -----ETGKPVTLKDVFGAYSDMTSTSTSGVNIIDSLNNPQD----- 201
Qy 237 PTLDSFKAIM-WDFV-PYFRMTK-RRHEIPTQGLAV---SRRVGIEMEQKQAVLGSA 290
Db 202 PFVENTKLLRFDFLDPFFLSITVFFFLPILEVLNICVFPREVTFNLRKSVKR-----M 256
Qy 291 SDQAVDKVQVQGRDILSLVLRANIANLPESQKLSDEEVLAQISNLLFAGYETSTVLTW 350

```

```

Db 257 KESRLEDTQKRVDFLQMDISQSKETESHKALSDELVAQSIIFIFAGYETTSVLSF 316
Qy 351 MEHRLSEDAVQKLRKEICQI--DTDMPTLDLNLALPYLEAFVKESLRLDPPSPYANRE 408
Db 317 IMVELATHPDVQKQLEIDAVLPNKAPPTDYVLQMEVLDVWVNETLELFFIAMRLERV 376
Qy 409 CLKDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDAEERPERWL 468
Db 377 CKKQVE-----INGMFIPKGVVMIPSYALHRDP-KYWTEPEKFLPERFS 420
Qy 469 EDVTDLSNS-IEAPYGHQASFIQGRACFGHFAVAEMKAEFLVTLRRVQREP 520
Db 421 KKNKDNIDPIYITPFG-----SGPRNCIGMRFPALMMKMLALIRVLQNFSPRP 467

RESULT 5
US-09-957-997-3
; Sequence 3, Application US/09957997
; Patent No. US20020150915A1
; GENERAL INFORMATION:
; APPLICANT: Berkenstam, Anders
; APPLICANT: Bertilsson, Gran
; APPLICANT: Blomquist, Patrik
; TITLE OF INVENTION: PROMOTER SEQUENCES
; FILE REFERENCE: 13425-046001
; CURRENT APPLICATION NUMBER: US/09/957,997
; CURRENT FILING DATE: 2001-09-21
; EARLIER APPLICATION NUMBER: SE 0003393-6
; EARLIER FILING DATE: 2000-09-22
; EARLIER APPLICATION NUMBER: 60/238,895
; EARLIER FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-957-997-3

Query Match
Best Local Similarity 29.0%; Score 385.5; DB 9; Length 503;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

Qy 13 LAAPSW--ASTAFPSLYL-APRRSLLYNLQ--PNHTNY--FTGNFLDILSARTG-----E 61
Db 7 LAMETWLLAVSLVLLYLYGTHSHGLFKLGIPOGTPLEFLGN---ALSFRKGYWTFDME 63
Qy 62 EHAKYREKYGSTLRFAGIAGAPVLSNSTDPKVFNHVM-KEAYD---YKPGGVAARVL 113
Db 64 CYKKYKRVNG-----IYDQQPMLAITDPDMIKTVLKECYSVFTNRRPFGPVGFMKNAI 118
Qy 114 RIATGDGVVTAEGEAHRRHRIIMPSLSAQAVKSMVPIFLKGMELVDKQMEDAAEKDMA 173
Db 119 SI-----AEDEWKRIRLSLSPTFTSGKLKEMVPIIAQYGDVLRNLRREA----- 164
Qy 174 VGESAGEKATRLLET-EGVDVQKWDVGRATLDVMAAGFDYKSDSLQNKTNELYVAFVGLT 232
Db 165 -----ETGKPVTLKHFVFGAYSDMTSTSTSGVSIIDSLNNPQD-----PFVENT 207
Qy 233 D---GFAPTLDSFKAIMWDFV-PYFRMTKSRHEIPTQGLAVS---RRVGIEMEQKQAV 286
Db 208 KLLLRFPN-LDPFVLSIKVP-PFL-----TPILEALNITVFPKRVISFTKSVQKIK 257
Qy 287 LGSASDAQVQKQVQGR--DILSLVLRANIANLPESQKLSDEEVLAQISNLLFAGYETS 344
Db 258 EGRL-----KETQKRVDFLQMDISQSKDSETHKALSDELMAQSIIFIFAGYETT 310
Qy 345 STVLTWVHRLSEDAVQKLRKEICQIDTM-----PTLDLNLALPYLEAFVKESLRLD 399
Db 311 SSVLSFIYELATHPDVQKQVQKVE---IDTVLENKAPPTDYVLQLEYLDVWVNETLRLF 367
Qy 400 PPSPYANRECLKQEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRSKFIYGEDA 459

```

Db 368 PVAMRLERVCKDVE-----INGMIPKGVVMIPSPVYLHDP-KYWTPE 411
Qy 460 BEFPERWLEDVTDLSNS-IEAPYGHQASFISSGPRACFGWPAVAEMKAFVTLRRVQF 518
Db 412 EKLPFRFSKGNKNDIPYTPFG-----SGPRNCIGMRFALVNMKLALVRVQNF 465
Qy 519 EP 520
Db 466 KP 467

RESULT 6
US-10-425-114-38180
; Sequence 38180, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38180
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-017-B5_FLI.ppep
US-10-425-114-38180

Query Match 13.5%; Score 384; DB 12; Length 537;
Best Local Similarity 27.4%; Pred. No. 1.4e-25;
Matches 162; Conservative 79; Mismatches 230; Indels 120; Gaps 24;
Qy 4 LVLLTGALGAFAFSWASIAFFS-LYAPRR-SSLNLOQPNHTNY--FTGNFLDILSART 59
Db 23 LLIVYGVIG-ALLMKAARLLDRLWEPRLERLALRAQGLRGTSYRFLTG---DLREYR 78
Qy 60 GBHAKYRE-----KYGSTLRFAGIAGAPVLNSTDPKVFNVVK 98
Db 79 SKBEAWARPLRCHDIAGHVEFFIHGAVLEHGKTC-FWFGVPVPRVTVPDLDARDMA 137
Qy 99 EAY-DYKPGMAARVLRATGCGVVTAEAGEAHRHRRIMIPSLSAQAVKSMVPIFLEKGM 157
Db 138 NKFGHEKPKFPA--LTKLFSDGVANHEGKWKVTHRRILNPAFHLEKLMPLAFSACCE 195
Qy 158 ELVDKMDAAEKDMAGVESAGEKATRLTEGVDVKGWGRATLDVMALAGF-----D 211
Db 196 ELVSRW-----AOSLPGDSCSEL-----DVPDELQTLTGDIVISRTAFGSSYLEG 239
Qy 212 YKSDSLQNKTNELVAFVGLTDFGAPTLDSFKAIMWDFVPYFRTMKRHEIPLTQGLAVS 271
Db 240 RKIFQLOAEQAERLMSII---DKFA-----VPGWMSLPTKNNRMQRKSEI 283
Qy 272 RRVGIELMEQKQAVLGASDAQVDKDVQGRDILSLVRANIANLPES-----QKLSDE 327
Db 284 DSILRGLIGKRMQAMKQGESD-----KD-----DLGLLESNARETGDSQSQPGGLTME 334
Qy 328 EVLAQISNLLFAGYETSSTVLTWMFHLSEDKAVQDKLEEICQI--DTPDPTLDLNLAL 385
Db 335 EVMECKLFPAGNETTSVLTWVLLSWHPWQDRAEEVLVFGKQKQGYDGLSEL 394
Qy 386 PYLEAFVKESRLRDPSPYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMPLN 445
Db 395 KTVTMILYEYLRYPYPAIAFSSRKYKE-----MVGVDVTPYAGVTLPELV 439

Qy 446 FNINRSKFIYGEDAEFRPERWLEDVTDLSNLSIEA--PYGHQASFISSGPRACFGWPAVA 503
Db 440 LFIHDDPDWSDAHERPFRPAEGVARASKDRLAFFPCW-----GPRICIGQNFALL 493
Qy 504 EKMAFLFVTLRRVQFEPHSHPEYEHHTLIISRRP--IVGREKEGYQMRLQ 552
Db 494 EAKWALSMLQRFQFE---LAPTPTVH-----PRRVIMLRPMHGAQIKLR 535

RESULT 7
US-10-425-114-67055
; Sequence 67055, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67055
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4757-032-B12_FLI.ppep
US-10-425-114-67055

Query Match 13.2%; Score 377.5; DB 12; Length 527;
Best Local Similarity 24.9%; Pred. No. 5.3e-25;
Matches 141; Conservative 106; Mismatches 212; Indels 107; Gaps 25;
Qy 6 LITGALGLAFAFSWASIAFFS---YLAPRRSSLNLOQPNHTNYFTGNFLDILSARTG-- 60
Db 24 LLIGAL-----FFLVKPYVTVTRWFRGOGIGPSY-REFVGLSEPEIKRMAAGS 71
Qy 61 -----EBHAKYREKYGSTURFAGIAGAPVLNSTDPKVFNVKWEAYD-Y 103
Db 72 KITLDVGDHDFVPIVQPYRWRVSDYKTFLY-WFGAVPTICVAEVLGVRQVLAERTGLF 130
Qy 104 KPGMAARVLRATGCGVVTAEAGEAHRHRRIMIPSLSAQAVKSMVPIFLEKGMELVDKM 163
Db 131 PKDYNDSS-MEVLLGKGLVLANGEDWKRHEVHPAFKPKDKLTKMSVVM-----ADLVQM 185
Qy 164 MEDAAEKDMAGVESAGEKATRLTEGVDVKGWGRATLDVMALAGF--DYKSDSLQNK 221
Db 186 MQQWRSQ-----IQRASNHEAE-IELSSSEFSELTSDVIAHTAFGTSYK-----EG 229
Qy 222 NELVAFVGLTDFGAPTLDSFKAIMWDFVPYFRTMKRHEIPLTQGLAVSRVRGIELMEQ 281
Db 230 KKVFFVA-----QKELQELTFSTWIDIPAPAC-RKLKLPSTKS---SRVV--EELDK 275
Qy 282 KQAVLGASDAQVDKDVQ--GRDILSLVRANIANLPESQKLSDEVLQISNLLFA 339
Db 276 KYRSLMLAIEGRLLAARGTSGVNDLLGLMLQAR-ALEQGHQMLTTEIIVDECKTFIPA 334
Qy 340 GYETSTVLTWMFHLSEDKAVQDKLEEICQIDTD-MPTLDLNLALPYLEAFVKESLRL 398
Db 335 GQDTTSHLLTWTMFLLSRYSEWQRLREVURECGDAVNPDPVTVKULNMVLLSRL 394
Qy 399 DPPSPYANRECLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMPLNFRNRSKFIYGED 458
Db 395 YSPVVFIR-----AVGSD--ILLSTRVPKGTMLSIPIALLHRDKDVGQD 439
Qy 459 ABEFPERWLEDVTDLSL-----NSIEAPYGHQASFISSGPRACFGWPAVAEMKAFVTLR 514
Db 440 ABEFNPDRFEHGSNAAAKHPNAL-----LSFSQGPRACTIGONFANLEARIATIAMILQ 492

```

QY      515 RYQEE--PIISHPEYEHITLIISRPR 538
Db      493 RFSFELSNYVHAPXEAUTLM---PR 515

RESULT 8
US-10-425-114-65616
; Sequence 65616, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yinhua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65616
; LENGTH: 547
; TYPE: PR1
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700100628_FLI.pep
US-10-425-114-65616

Query Match      13.2%; Score 376.5; DB 12; Length 547;
Best Local Similarity 24.3%; Pred: No.6.9e-25;
Matches 144; Conservative 96; Mismatches 220; Indels 133; Gaps 25;

QY      7 LTGALGIAFSW--ASIAFFSFLYAPRSSLYNQPNHNY--FTCNFLDILISARTGEE 62
Db      37 LAGAVASVSLMLVAMTLEWAWMTPLRLDRALRAQLKGTRYRLFTGDLRE--TARVNR 94
QY      63 -----HAKYREKYGSTLRFAGIAGAPVLNSTPKVFNVHMKEAYD 102
Db      95 ARKKPLPLGCHDITPRVQPMHSHSTIKYKG-LSFTWFGTPTRVMIPODELKVELSNKFG 153
QY      103 Y--PKPGMAARVIRIATGDGVVTAEGEBAHKRHRIRIMTSPLSAQAVKSMVPIFLEKGMEL 159
Db      154 HFGKPRSSRIGRL-----ANGLVNHGGEKAWKRRILNPAFHHEKIKGMWVFSTCCIE 209
QY      160 VDKMDEDAEKONAVGESAGEKATLETEGVYKDWVGRATLDVMALAGF--DYKSD-- 215
Db      210 ITRW-----DNMS-SEGSE-----IDVMPEFQNTGDIISRTAFGNYQGR 253
QY      216 --SLQKTNELVYAFVGLTGDGAPTLDSPKALMDFVP--YFRWTKGRHEIPLTQGLAVS 271
Db      254 IPELOELAERLI-----QSVQTFIFGYWFLPKNNRM----- 288
QY      272 RRVGIELMEOKQAVLGASDAQVKQVGR-----DILSLVRAI--AANLPESOKLS 325
Db      289 RAIDVEIRKILR-IIGKR-----EKDTKRNKTDLLGLLESNTROSGNASLGLT 341
QY      326 DEEVLAIQISNLLFAGYETSTVLTMWPHRLSDKAVQOKLREEI--CQIDTDMPTLDNEL 384
Db      342 TEDVIECKLFYFAGNETTSVLLTWTLIVLSMHPWEQERAREEVLSHFGRTTPDYDSL 401
QY      385 LPYLEAPVKESRLDPPSPYANRECKDBDFTPLAEPVIGRDGGSVINEVRITKWTWMLP 444
Db      402 LKVTMTLHEVLVLYPLATFLTRTYKEME-----LGGIKYPAGVEILLP 446
QY      445 LPINRSKFTYGDAAEEFERPWLVEDVTDLSNLSIAPYGHQASFI---SGPRACFGWRFA 501
Db      447 VIFLHDDPDILWGDAEFTNPERF-----ANGISATRHQAAFPFGGPRICIGQSPA 499
QY      502 VAEMKAFELVTLRRVOFE--PIISHPEYSHITLIISRPRIVGREKBYQMQLQ 552

```

```
RESULT 10
US-10-425-114-59349
; Sequence 59349, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59349
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-OSUELIB3474019G01_FLI.pep
US-10-425-114-59349

Query Match      12.9%; Score 368.5; DB 12; Length 560;
Best Local Similarity 24.6%; Pred. No. 3.8e-24;
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLTGALGALAFWSIASIAFFS-LYLAPRR-SSLYNLQGNHNY--FTGNF----- 51
Db 49 ILAAAAAVALLLAVSTLEWAWTPRRLERALLRAQGIIRNRYRLFTGDPVENVRLNR 108
QY 52 -----LDLS-----ARTGEHAKYREKYSTLRERAGIAGAPVLNSTDPKVF 94
Db 109 EARKPLPLGCHDIIIPRVLPMSKAVEHGK-----PSTWFGTPRVMISDPSIR 160
QY 95 HVNKEA---YDYPKPGMAARVLRIATGCVVTAEGEAHKKRRRIMIPSLSAQAVKSMVPI 151
Db 161 EVMSNKGFGYKPKTRLGKLL---ASGVSYEGEKWAKHRRILNPAFHHEKIKRMLPV 216
QY 152 FLEKGMELVDKMDAEDAEKDMVAGESAGEKATRETEGVVDKVGVRATLDVWALAGF- 210
Db 217 FSNCTEMVTRW-----ENSMIS-EGMSE-----VDVWPEFQNLTDGVISKTAFG 260
QY 211 -DYKSD----SLQKTNELNYVAFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPT 265
Db 261 SSYEGRIRIFQQAESAERII-----QAFRTI---FIPGYWFLPTKNNRRLR 304
QY 266 QGLAVSRVGIELMEQKQAVLGSASDQAVDKQVQGRDILSLVRANI-AANLPESQKL 324
Db 305 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLLGLLVESNMRESNGKAE LGM 355
QY 325 SDEEVLQIISNLLFAGYETSSVLTWTFHRLSDKAVQDKLREIC-QIDTDMPTLDELN 383
Db 356 TTDEIIECKLFYFAGMETTSVLLTWLTVLSMHPWQERAREEVLHFGRTTPDYDSL 415
QY 384 ALPYLEAFVKESLRDPPSPYANRECKDBDFPLAEPIVIGRDSVINEVRITKTMVWL 443
Db 416 RLKIVTMILYEVRLYPPVFLTRRTYKEMEL-----GGIKYPAEVT-----LML 460
QY 444 PLFNINRSKFIYGEDAESEFRPERWLEDVTDLSNIEAPYGHQASFS---GPRACFGWRF 500
Db 461 PILFIHDDPDINGKDAEFPNFGFADGISNATK-----YQTSFPFGWGPRICIGQNF 513
QY 501 AVAEMKAFVTLRRVQFEPIISHPEYEH--ITLIISRRPRIVGREKEGYQMLQ 552
Db 514 ALLEAKQAICTILQRFSE---LSPSYIHAPFTVITLHP-----QHGAQIKLK 558
```

```
RESULT 11
US-10-425-114-59350
; Sequence 59350, Application US/10425114
```

```
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59350
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-OSFLCYP143F07_FLI.pep
US-10-425-114-59350
```

```
Query Match      12.9%; Score 368.5; DB 12; Length 562;
Best Local Similarity 24.6%; Pred. No. 3.8e-24;
Matches 146; Conservative 93; Mismatches 226; Indels 129; Gaps 27;

QY 4 LVLTGALGALAFWSIASIAFFS-LYLAPRR-SSLYNLQGNHNY--FTGNF----- 51
Db 51 ILAAAAAVALLLAVSTLEWAWTPRRLERALLRAQGIIRNRYRLFTGDPVENVRLNR 110
QY 52 -----LDLS-----ARTGEHAKYREKYSTLRERAGIAGAPVLNSTDPKVF 94
Db 111 EARKPLPLGCHDIIIPRVLPMSKAVEHGK-----PSTWFGTPRVMISDPSIR 162
QY 95 HVNKEA---YDYPKPGMAARVLRIATGCVVTAEGEAHKKRRRIMIPSLSAQAVKSMVPI 151
Db 163 EVMSNKGFGYKPKTRLGKLL---ASGVSYEGEKWAKHRRILNPAFHHEKIKRMLPV 218
QY 152 FLEKGMELVDKMDAEDAEKDMVAGESAGEKATRETEGVVDKVGVRATLDVWALAGF- 210
Db 219 FSNCTEMVTRW-----ENSMIS-EGMSE-----VDVWPEFQNLTDGVISKTAFG 262
QY 211 -DYKSD----SLQKTNELNYVAFVGLTDGFAPTLDSFKAIMWDFVPYFRTMKRHEIPT 265
Db 263 SSYEGRIRIFQQAESAERII-----QAFRTI---FIPGYWFLPTKNNRRLR 306
QY 266 QGLAVSRVGIELMEQKQAVLGSASDQAVDKQVQGRDILSLVRANI-AANLPESQKL 324
Db 307 E---IEREVSKLL-----RGIIGK-RERAIKNGETSNGLLGLLVESNMRESNGKAE LGM 357
QY 325 SDEEVLQIISNLLFAGYETSSVLTWTFHRLSDKAVQDKLREIC-QIDTDMPTLDELN 383
Db 358 TTDEIIECKLFYFAGMETTSVLLTWLTVLSMHPWQERAREEVLHFGRTTPDYDSL 417
QY 384 ALPYLEAFVKESLRDPPSPYANRECKDBDFPLAEPIVIGRDSVINEVRITKTMVWL 443
Db 418 RLKIVTMILYEVRLYPPVFLTRRTYKEMEL-----GGIKYPAEVT-----LML 462
QY 444 PLFNINRSKFIYGEDAESEFRPERWLEDVTDLSNIEAPYGHQASFS---GPRACFGWRF 500
Db 463 PILFIHDDPDINGKDAEFPNFGFADGISNATK-----YQTSFPFGWGPRICIGQNF 515
QY 501 AVAEMKAFVTLRRVQFEPIISHPEYEH--ITLIISRRPRIVGREKEGYQMLQ 552
Db 516 ALLEAKQAICTILQRFSE---LSPSYIHAPFTVITLHP-----QHGAQIKLK 560
```

```
RESULT 12
US-09-992-901-2
; Sequence 2, Application US/09992901
; Patent No. US2002007346A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M.
```

APPLICANT: Chory, Joanne
TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
FILE REFERENCE: SALKINS.024DV1
CURRENT APPLICATION NUMBER: US/09/992,901
CURRENT FILING DATE: 2001-11-14
PRIORITY APPLICATION NUMBER: US 09/527,073
PRIORITY FILING DATE: 2000-03-16
PRIORITY APPLICATION NUMBER: US 60/124570
PRIORITY FILING DATE: 1999-03-16
PRIORITY APPLICATION NUMBER: US 60/170,931
PRIORITY FILING DATE: 1999-12-14
PRIORITY APPLICATION NUMBER: US 60/172,832
PRIORITY FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 520
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-992-901-2

Query Match 12.9%; Score 367; DB 9; Length 520;
Best Local Similarity 22.2%; Pred. No. 4.6e-24;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;
QY 4 LVLLTGALGAASFASIAFFSLYLAPR-----SSLYNLQGNHNTNYFTGNFLDLS--- 56
DB 13 VLVLSTVLIVVKGSL-----LWPRKIEBHFQKQIRGPY-HFPIGNVKGVLGVM 67
QY 57 -----ARTGEHAKRYEKYSTLRFAAGIAGAPVNST--DPKVFNNHVKMEA 100
DB 68 KASSHPMPFPHNLPVLSFYHWRKIYATFL--VWFGPFRLTADVADPOLIREIFS 124
QY 101 YDPKPGMAARVRIATGCVTTAECAHRRRIMIPSLSAQAVKSMVPIFEKGMELV 160
DB 125 EFYEK-NEAHPVLQLEGDLGSLKGEWAHRRKIISPTFHNNKLLVPPVVLKSVTDMV 183
QY 161 DKWMDAAEKDVAVGSAGEKATRLTEGVDYDQVGVGRATLDVMAAGFDYKSDSLQNK 220
DB 184 DKMSDKLSEN-----GEVE-----VDVYENFQILTEDVISRTAFGSSVE----- 222
QY 221 TNELYAVFGLTQGFAP-----TLDSFKAINMDVPPVFRMKRHEIPLTQGLAV 270
DB 223 -----DGRAVFLQAQMLLCAEAFQKV---FIPGYRFPF-----TRGNLK 260
QY 271 SRRVG-----TELMEQKQAVLGSASQAVDKQVQGRDILSLVRANIAANLPESQK 323
DB 261 SRKLDKEIRKSLKLTERRQNAIDGEGEC---KEPAKDLGLMIQA-----KN 308
QY 324 LSDEVLQAQSNLLFAGYETSTVLTWFMHRLSEDKAVQDKLREICQI--DTMPTLDE 381
DB 309 VTQDIVEECKSFPFAGKQTTLSLLTWTLLSMHPQAKARDEVLRVCGSRDVPYTKDH 368
QY 382 LNALPYLEAFKESLRDPPSPVANRECLNDEDFIPLABPVGDRGVSINEVRIKGTMV 441
DB 369 VVKUKTSLMINESLRLYPIVATIRAKSDVK-----LGGYKIFCGTEL 413
QY 442 MLPFNINRSKFYGEDAEFRPERMLEVDYDLSNIEAPYGHQASFTS---GPRACFGW 498
DB 414 LIPIIAVHHDQAIWNGDWNENFARFADGVPAAK-----HPVGFIPFGLGVRTCIGQ 466
QY 499 RFAVAEMKAFVTLRVRQVEPIISHPEYEHITLII 534
DB 467 NLAIQAKLTUAVMIQRTFH---LAPTYOHAPTVL 499

RESULT 13
US-10-425-114-38827
Sequence 38827, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 38827
LENGTH: 662
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700045236_FLI.pep
US-10-425-114-38827

Query Match 12.8%; Score 365.5; DB 12; Length 662;
Best Local Similarity 22.9%; Pred. No. 9.1e-24;
Matches 123; Conservative 96; Mismatches 193; Indels 125; Gaps 16;
QY 70 YGSTRFAGIAGAPVNSTDPKVFNNHVKMEAYDPKPGMAARVRIATGCVTTAECAH 129
DB 188 YGGIFRL-NFGPKSLIYSDPAIAKHILRENSKAYSIGILAEILFVNGTGLPADGEIW 246
QY 130 KHRHRIMIPSLSAQAVKSMVPIFEKGMELVDKMDAAEKDVAVGSAGEKATRLTE 189
DB 247 RYRRRAIIPALHOKVYTAIIGLGEASQRLCEKL-----DKAA--VDGE 288
QY 190 GVDVMDWGRATLDVMAAGFDYKSDSLQNTNELYAVFVGLTQGFAPTLDSFKAINWDF 249
DB 289 DMEMESLFSRLTLDVIGKAVFNYPDSLSYDNGIVEAVYVTLRE--AEMRSTSPITW-- 344
QY 250 VYFRTMKRHEIPLTQGLAVSRVGIEMQKQAVLGSASQD-----AVDKQDVQG 302
DB 345 -----EIPWKDISPRQKVNALK-----LINSTLDELIAICKLVEQEDLQF 388
QY 303 RD-----ILSLVRANIAANLPESQKLSDEVLQAQSNLLFAGYETSTVLTWFMH 353
DB 389 HBEYMEQDPSILHFL-----LASGDDVSSKQLRDLMTMLIAGHETSAAVLTWTFY 440
QY 354 RLSDEKAVQDKLREICQIDTD-MPTLDLALPYLEAFVKESLRDPPSPYANRECLKD 412
DB 441 LSKYKPVNAKQDRAQSVLGGDLFTIEDVKLKYTRVINESRLYPQPVLIERSLED 500
QY 413 EDFIPLAEFVIGRDSVINEVRIKGTVMMLPLFNINRSKFYGEDAEFRPERMLEVDVT 472
DB 501 D-----ILGGYPIGRGEDIFISVNNLHCP-KHWDAAEVFNPERWPLDGP 544
QY 473 DLSNIEAPYGHQASFTSGPRACFGWRFVAEMKAFVTLRVRQVE----- 519
DB 545 NP-NEINQFNSY-LPFGGGRKCVGDMFATFETVVTATMLVKRFDQFMAPGAPPVDMTGT 602
QY 520 -----PIISHPEYEHIT-----LIISRPIVGRKEG 546
DB 603 ATHTTEGLKMTVTRTPPVPINLEMKIITDSQBSTLSAPSMVVSASVASGEDQG 659

RESULT 14
US-10-114-270-60
Sequence 60, Application US/10114270
Publication No. US20040030110A1
GENERAL INFORMATION:
APPLICANT: Guo, Xiaojia
APPLICANT: Kekuda, Ramesh
APPLICANT: Miller, Charles E.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Spyttek, Kimberly A.
APPLICANT: Patturajan, Meera
APPLICANT: Liu, Ziaohong
APPLICANT: Gusev, Vladimir Y.

[illegible]

Qy	268	LAVSRVGIEMEQKQOAVLGSASQADVQKDVQGRDILSLVRA-----NIA	315
Db	265	---DRVIKASLTDMIKK-----REKAPKTGEATRDDLLGILLESNHKEIOEHRNNENVG	315
Qy	316	ANLPESQKLSDEVLQAQISNLLFAGYETSTVLTWMFHLSEDKAVQDKLREEICQI-DT	374
Db	316	MNL-----NDVIECKLFYFAGQETTSVLLVWTVLLSRYPDWQSRAREEVLQVFGK	367
Qy	375	DMPTLDELNALPYLE-AFVKESLRLDPSPYANRECKDEDEFIPLAEPVIGDGSVINEV	433
Db	368	QAPNFDGLSHKIWTMIFLKVVLRLYPVAVGLNRN-----VDRDMKLG-N-L	412
Qy	434	RITKGTWMLPLFNINRSKFIYGEDAEERPERWLEDVTDLSN:IEA--PYGHQASFISG	491
Db	413	SLPAGVQVSLPTTMVPHDRELWGDVDNEFKPERFSEGVLKATNGRVSPFPFCM-----G	466
Qy	492	PRACFGWRFAVAENKAPLFTVTLRRVQFEPFIISHPEYEH--ITLIISRPR	538
Db	467	PRICIGQNFSLLEAKMALSTILOHFSPE---LSPAYAHAPVTFTLQPO	512

Search completed: April 2, 2004, 14:10:37
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 17 Seconds

(without alignments)
3151.687 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852

Sequence: 1 MFILVLLTGALGLAAFSWAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78: *
1: P1r1: *
2: P1r2: *
3: P1r3: *
4: P1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	419	14.7	593	2 F86441	probable cytochrom
2	406	14.2	503	2 JC4702	cytochrome P450 3A
3	404	14.2	503	2 S14275	steroid 6beta-mono
4	402	14.1	501	2 A34236	cytochrome P450 3A
5	400	14.0	526	2 T02191	cytochrome P450 ho
6	399	14.0	504	2 A29410	cytochrome P450, S
7	398.5	14.0	503	2 A29815	cytochrome P450 3A
8	398	14.0	503	2 S50211	cytochrome P450 3A
9	395.5	13.9	503	2 S28168	nifedipine oxidase
10	394	13.8	504	2 A60564	cytochrome P450 3A
11	385.5	13.5	503	2 JX0062	cytochrome P450 3A
12	385	13.5	502	1 A34101	cytochrome P450 3A
13	379	13.3	501	2 A29487	cytochrome P450 3A
14	370	13.0	504	2 A22631	cytochrome P450 3A
15	368	12.9	504	2 S50892	cytochrome P450 3A
16	365	12.8	520	2 H84663	probable cytochrom
17	363.5	12.7	511	2 T00864	cytochrome P450 ho
18	362.5	12.7	523	2 B96662	probable cytochrom
19	359.5	12.6	511	2 S66472	cytochrome P450 4B
20	357	12.5	502	2 JX0334	cytochrome P450 3A
21	355.5	12.5	512	2 A96695	hypothetical prote
22	350.5	12.3	1054	1 A49975	NADPH-ferrihemopro
23	350	12.3	526	2 JC4533	cytochrome P450 4F
24	349.5	12.3	511	1 O4HUB1	cytochrome P450 4B
25	349.5	12.3	524	2 S29723	cytochrome P450 4F
26	346.5	12.1	511	1 B40164	cytochrome P450 4B
27	346.5	12.1	524	2 JC7594	cytochrome P450 en
28	346.5	12.1	524	2 JC7598	cytochrome P450 en
29	346	12.1	537	2 T02450	probable cytochrom

30 345.5 12.1 520 2 T24778
31 345 12.1 516 2 T00514
32 344 12.1 520 1 A46661
33 343 12.0 504 2 A25222
34 340.5 11.9 524 2 T09944
35 340 11.9 517 2 T02192
36 339.5 11.9 508 2 T16980
37 338.5 11.9 503 2 JC7627
38 337.5 11.8 524 2 T09999
39 337 11.8 520 2 S45702
40 336.5 11.8 490 2 T00404
41 336 11.8 522 2 JC4532
42 334.5 11.7 509 2 T24785
43 334.5 11.7 576 1 H71414
44 333.5 11.7 497 2 S57097
45 333.5 11.7 499 2 T18699

ALIGNMENTS

RESULT 1

P86441

probable cytochrome P450 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001

C/Accession: P86441

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: P86441

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-593 <STC>

A/Cross-references: GB:AE005172; NID:g1136728; PIDN:AAG31309.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology

C/Keywords: heme; iron; metalloprotein

F,514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match

Best Local Similarity 14.7%; Score 419; DB 2; Length 593;

Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NFDLLSATG--EHHAKYREKVGST-----LRFAGI-----AGAPVLN 86
DB 93 NVLDFFWDTGSDQYPKVPEAKGSIQAVRNEAFIPLYELFLTYGGIFLTGPKSFLI 152
QY 87 STPKVFNHMK--AYDPKPGMAARVLIATGCVTAGEAHKRRHRIIMPISLSAQV 145
DB 153 VSDPSIAKHILKDNKAYSK-GILAEILDVFMGKLIPADGEIWRRRRAIVPALHOKYV 211
QY 146 KSMVPFLEKGMELVDKMMEDAAEKMAVGESAGEKKATRLTEGVVDKDWVGRATLDVM 205
DB 212 AMISLFGASDELQCKL--DAA-----ALKGEVEMESLPSRLTLDII 253
QY 206 ALAGFYKSDSLQNTNELYVAVGLTGDGAPTLDLSFKAIMDVPVFRFWKRRHEIPLT 265
DB 254 GKAVFYDFDSLNTDNTGVIEAVYTVLREADRSPSP--VWD-IPIWKDIS-----PRQ 305
QY 266 QGLAVRRRGIELMEOKKQAVLGSASDAVKKDVQ-----GRD--ILSLVRANIAA 316
DB 306 RKVATSLKINTLDD-----LIATCKMWEESLQFHEEYMERDPSILHFL----- 353

RESULT 6

A29410
cytochrome P450, glucocorticoid-inducible, hepatic - human
N:Alternate names: cytochrome P450 3A3; cytochrome P450 HLP
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1989 #sequence revision 31-Mar-1989 #text_change 28-Jul-2000
C:Accession: A29410; A25170; A25450
R:Motowa, D.T.; Schuetz, E.G.; Wrighton, S.A.; Watkins, P.B.; Kremers, P.; Mendez-Picon, G.
Proc. Natl. Acad. Sci. U.S.A. 83, 5311-5315, 1986
A:Title: Complete cDNA sequence of a cytochrome P-450 inducible by glucocorticoids in human liver.
A:Reference number: A29411; MUID:86259780; PMID:3460094
A:Accession: A29410
A:Molecule type: mRNA
A:Residues: 1-504 <MO>
A:Cross-references: GB:M13785; NID:G181353; PIDN:AAA35742.1; PID:G181354
R:Watkins, P.B.; Wrighton, S.A.; Maurel, P.; Schuetz, E.G.; Mendez-Picon, G.; Parker, G.
Proc. Natl. Acad. Sci. U.S.A. 82, 6310-6314, 1985
A:Title: Identification of an inducible form of cytochrome P-450 in human liver.
A:Reference number: A25170; MUID:85298342; PMID:3898085
A:Accession: A25170
A:Molecule type: protein
A:Residues: 2-21 <WAT>
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; liver; metalloprotein; monooxygenase
F:303-465/Domain: cytochrome P450 homology <P45>
F:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 399; DB 2; Length 504;
Best Local Similarity 28.4%; Pred. No. 1.9e+18;
Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;

QY 13 LAAFSW--ASTAFSLYLAPRRS-SLYNLOG-PNHNY-FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLAVSLVLLYGYTHSHGLFKKLGIPGTPPLFLGN---ILSYHKGFCMFDM 63

QY 62 EHAKYREKYGTLRFAGTAGAPVLNSTDPKVFNVHM-KEAYD-----YKPGMAARVL 113
DB 64 CHKYGKVMGF--YDG-QQPVLAITDPDMIKLVLVKCYSVFTNREFPGVPMKSAI 118

QY 114 RIATGDGVVTAEGAHKRRHIMPSLSAQAVKSMVPIFEKGMELVDKMWEDAAEKDA 173
DB 119 SI-----AEDEWKRLRLSLSPFTSGKLEKEMVPIIAQYGDVLVRLRE----- 163

QY 174 VQESAGEKATRLTEGVYDKDWGRATLDVMAAGFDYKSDSLQNKNELYVAFVGLTD 233
DB 164 -----RETGKPVTLKDFGAYSMVDVITSSFGVNVDSLNNPD----- 201

QY 234 GFAPTLDSFKAIM-WDFV-PYFRTMK-RREIPLTQGL-----AVSRVG 275
DB 202 ---PLVENTKLLRFDLDFDFUSITVFPFLIPILVLMICVPPREVTNFKAVKRMKE 258

QY 276 TELMEKQKQAVLGASDAQVDKQDVQGRDILSLVRAANTAAALPSSQK-LSDEEVLQA 334
DB 259 SRLDETQHRV-----DFLQLMIDSHKNSKETESHKALSLELVAQSI 301

QY 335 NLLFAGYETSTVLTWPHRLSDEKAVODKLRREICQI---DTEMPTLELNAFLVLEAFV 392
DB 302 IPFAGYETTSVLVSFIMYELATHPDVQKQLQBEIDAVLPNKAPPTDVLQMEYLDWV 361

QY 393 KESRLDPPSYANRECLKQEDFIPLAEPVIGRGGVINEVRITKGTVMVLPFLFNRSK 452
DB 362 NETLRLFPIMRLERLVCKDVE-----INGMFTPKGVVMIPTSYALHRDP 406

QY 453 FYGEDAEERPERMLEDVTDLSN-TEAPYGHQASFISSPRACQWRFVAVKAFV 511
DB 407 -KNTWEPSEKFLPERSEKQKNDNDPIYITPFG-----SGPRNCIGMRFALNMKALIR 459

QY 512 TLRRVQFER 520
DB 460 VLQNFSEKFP 468

RESULT 7

A29815
cytochrome P450 3A4 nifedipine oxidase (EC 1.14.14.-) - human
N:Alternate names: cytochrome P450 (PCN1); cytochrome P450-HM1
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1989 #sequence revision 30-Sep-1989 #text_change 28-Jul-2000
C:Accession: A29815; A25517; A32199; PX0012; S03851
R:Gonzalez, F.J.; Schmid, B.J.; Umeno, M.; McBride, O.W.; Hardwick, J.P.; Meyer, U.A.; (DNA 7, 79-86, 1988
A:Title: Human P450PCN1: sequence, chromosome localization, and direct evidence through
A:Reference number: A29815; MUID:88195781; PMID:3267210
A:Accession: A29815
A:Molecule type: mRNA
A:Residues: 1-503 <GON>
A:Cross-references: GB:M18907; NID:G181373; PIDN:AAA35745.1; PID:G181374
R:Spurr, N.K.; Gough, A.C.; Stevenson, K.; Wolf, C.R.
Hum. Genet. 81, 171-174, 1989
A:Title: The human cytochrome P450 CYP3 locus: assignment to chromosome 7q22-qter.
A:Reference number: S16900; MUID:89108438; PMID:2563251
A:Accession: S16900
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-391 'W' 393-503 <SP2>
A:Cross-references: EMBL:X12387; NID:G35910; PIDN:CAA30944.1; PID:G35911
R:Beaune, P.H.; Umbenhauer, D.R.; Bork, R.W.; Lloyd, R.S.; Guengerich, F.P.
Proc. Natl. Acad. Sci. U.S.A. 83, 8064-8068, 1986
A:Title: Isolation and sequence determination of a cDNA clone related to human cytochrome
A:Reference number: A25517; MUID:87041402; PMID:3464943
A:Accession: A25517
A:Molecule type: mRNA
A:Residues: 1-391 'W' 393-503 <BEA>
R:Bork, R.W.; Muto, T.; Beaune, P.H.; Srivastava, P.K.; Lloyd, R.S.; Guengerich, F.P.
J. Biol. Chem. 264, 910-919, 1989
A:Title: Characterization of mRNA species related to human liver cytochrome P-450 nifedipine
A:Reference number: A32199; MUID:89093163; PMID:2463251
A:Accession: A32199
A:Molecule type: protein
A:Residues: 'X', 2-9 'XXX', 13-15 <BOR>
R:Komori, M.; Hashizume, T.; Ohi, H.; Miura, T.; Kitada, M.; Nagashima, K.; Kamataki, T.
J. Biochem. 104, 912-916, 1988
A:Title: Cytochrome P-450 in human liver microsomes: high-performance liquid chromatogr
A:Reference number: PX0012; MUID:89214010; PMID:3243766
A:Accession: PX0012
A:Molecule type: protein
A:Residues: 'X', 2-11 'X', 13-25 <KOM>
A:Experimental source: liver microsomes
C:Genetics:
A:Gene: GDB:CYP3A4
A:Cross-references: GDB:118782
A:Map position: 7q22.1-7q22.1
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metall
F:302-464/Domain: cytochrome P450 homology <P45>
F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 398.5; DB 2; Length 503;
Best Local Similarity 28.7%; Pred. No. 2e+18;
Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;

QY 13 LAAFSW--ASTAFSLYLAPRRS-SLYNLOG-PNHNY-FTGNFLDILSARTG-----E 61
DB 7 LAMETWLLAVSLVLLYGYTHSHGLFKKLGIPGTPPLFLGN---ILSYHKGFCMFDM 63

QY 62 EHAKYREKYGTLRFAGTAGAPVLNSTDPKVFNVHM-KEAYD---PKPGMAARVLRIAT 117
DB 64 CHKYGKVMGF--YDG-QQPVLAITDPDMIKLVLVKCYSVFTNREFPGVPMKSA- 117

QY 118 GDGVVTAEGAHKRRHIMIPSLSAQAVKSMVPIFEKGMELVDKMWEDAAEKDAVGES 177
DB 118 ---ISIADEWKRLRLSLSPFTSGKLEKEMVPIIAQYGDVLVRLRE----- 164

178 AGEKATRIET-EGVDVQKDWGATLDVWALAGFYKSDSLQNTKTNLYVAFVGLTDGFA 236
 165 -----ETGKFTVTKDVGAYSMVDVITSTSFQVNDISLNNPD----- 201
 237 PTLDSFKAM-WDFV-PYERTMK-RRHEPQLQGLAV---SRRVGIELMEOKQOAVLGS 290
 202 PFVENTKLLRFPDPLPFLSTVFPFLPILLEVNLNCPVPREVTNLRKSVKR-----M 256
 291 SDQAVKQVQGRDILSLVRANIANLPSOKLSDBEVLQAI SNLLPAGVETSTVLTW 350
 257 KESRLEDTQGRVDFLQMLIDQNSKSTESHKALSDELVAQSIIFIPAGVETSTVLSF 316
 351 MFRHLSKAVQDKREEICQI--DMDPTLDELNALPYLEAFVKSRLRDLPPSPYANRE 408
 317 IMTELATHPDVQKQLEEDAVLPNKAPTYDTVLMQMEYLDVWVNETLRLPIAMRLERV 376
 409 CLKDEDFIPLAEPVIGRDSGVINEVRITKGTVMPLFNINRSKIYGEDAEERFARWL 468
 377 CKXQVE-----INGMFIPKGVVVMIPSYALHRDP-KYWEPEKFLPERFS 420
 469 EDVTDLSNS-IEAPYGHQASFGSPRACFGWFAVAKAFLVTLRRVQPEP 520
 421 KKKNDIDPIYTPFG-----SGPRNCIGMRPALMMKALIRVLQNFESFKP 467

RESULT 8
 S50211
 cytochrome P450 3A13 - mouse
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Mus musculus (house mouse)
 C>Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-Jul-2000
 C/Accession: S50211; S18155
 R:Yanagimoto, T.; Itoh, S.; Sawada, M.; Hashimoto, H.; Kamataki, T.
 Biochim. Biophys. Acta 1201, 405-410, 1994
 A>Title: Molecular cloning and functional expression of a mouse cytochrome P-450 (Cyp3a-4)
 A/Reference number: S50211; MUID:95101705; PMID:7803471
 A/Accession: S50211
 A/Molecule type: mRNA
 A/Residues: 1-503 <YAN>
 A/Cross-references: EMBL:X63023; NID:G50634; PIDN:CAA44754.1; PID:G50635
 C/Genetics:
 A/Gene: CYP3A13
 C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
 C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
 F:302-464/Domain: cytochrome P450 homology <P45>
 F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 14.0%; Score 398; DB 2; Length 503;
 Best Local Similarity 27.0%; Pred. No. 2.2e-18;
 Matches 142; Conservative 74; Mismatches 179; Indels 130; Gaps 20;
 63 HAKYREKYGSTLRPAGIAGAPVLNSTDPKVNHN-KEAYD-----YPKPGMAARVLR 114
 65 HKYKGMWG---LYDG--RQVLAITDPIIKTVLKECYSTFNRRRFGVGLKKALS 119
 115 IATGQGVVTAGEAKHRRIRIMIPSLAQAVKSMVPIFLEKGMELVDMKMDAAEDMAV 174
 120 I-----SENEEKRIRALLSPTFTSGRLKEMFPII-----NQFTDVLVNRQ 162
 175 GESAGEKATRIETEGVDVQKDWGATLDVWALAGFYKSDSLQNT----- 221
 163 G-LGEGKPT-----SMKDI FGAYSMVDVITATSFQVNDISLNNPDQPFVEIKKLLKF 213
 222 ---NELYVAFVGLTDGFAPTLDSFKAIW--DFVPYERTMKRREIPLTQGLAVSRVGI 276
 214 DIFDLPLUS-VTLFPFLTPVDLNVSLFPDRDVISFTT-----SVERKEN 259
 277 ELMEQKQAVLGSASDAQVKKDQVGRDILSLVRANIANLPSOKLSDBEVLQAI SNL 336
 260 RMKEKEKQV-----DFLQMINSONYKTKESHKALSDEVEIQAQSVIF 302
 337 LFAGVETSTVLTWPHRLSEDKAVODKLRBEICOLDTDMF-----TLDLNLALPYLEAF 391

303 IPAGYETTSSALSALYLLALHPDVKQLQDE---IDAALPNKAPATYDTLLQWEYLDNV 359
 392 VKESIRLPPPPYANRECLKDEDEFIPLAEPVIGRDSGVINEVRITKGTVMPLFNINRS 451
 360 VNETLRLPIAGRLERVCKTVE-----INGLFIKPGTVMIPFALHKD 404
 452 KFIYGEDAEERFARWLSDVTDLSNSIEAPYGHQASFGSPRACFGWFAVAKAFLV 511
 405 P-KYWEPEERFARPERFSKKNQDSIN---PYMY-LPFGSGPRNCIGMRPALINMKVALVR 459
 512 TLRRVQPEPI-----ISHREYEHITLIISRPVIGRE 543
 459 VLQNTVQPCKEITPLKSKQGLQLENPLLLKVSARDETVDSE 503

RESULT 9
 S28168
 nifedipine oxidase (EC 1.14.14.-) cytochrome P450 3A8 - crab-eating macaque
 N/Alternate names: cytochrome P450 CMLc; cytochrome P450 MKnf2; cytochrome P450-MK2
 C/Species: Macaca fascicularis (crab-eating macaque)
 C/Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 28-Jul-2000
 C/Accession: S28168; S04509; S36875
 R:Komori, M.; Kikuchi, O.; Sakuma, T.; Funaki, J.; Kitada, M.; Kamataki, T.
 Biochim. Biophys. Acta 1171, 141-146, 1992
 A>Title: Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity of the pri
 A/Reference number: S28168; MUID:93129612; PMID:1282830
 A/Accession: S28168
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-503 <KOM>
 A/Cross-references: GB:S53047; NID:G263692; PIDN:AA824952.1; PID:G263693
 R:Ohta, K.; Kitada, M.; Hashizume, T.; Komori, M.; Ohi, H.; Kamataki, T.
 Biochim. Biophys. Acta 996, 142-145, 1989
 A>Title: Purification and characterization of two forms of hepatic microsomal cytochrome
 A/Reference number: S04509; MUID:89287352; PMID:2500151
 A/Accession: S04509
 A/Molecule type: protein
 A/Residues: 1-22 <OHT>
 R:Omori, S.; Horie, T.; Guengerich, F.P.; Kiuchi, M.; Kitada, M.
 Arch. Biochem. Biophys. 305, 405-413, 1993
 A>Title: Purification and characterization of two forms of hepatic microsomal cytochrome
 A/Reference number: S36874; MUID:93364294; PMID:8373178
 A/Accession: S36875
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-13 <OHM>
 C/Genetics:
 A/Gene: CYP3A8
 C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
 C/Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
 F:302-464/Domain: cytochrome P450 homology <P45>
 F:442/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 13.9%; Score 395.5; DB 2; Length 503;
 Best Local Similarity 28.6%; Pred. No. 3.2e-18;
 Matches 155; Conservative 85; Mismatches 187; Indels 115; Gaps 26;
 13 LAAPSW--ASIAFPFSLYAPRRS-SLYNLQG-PNHNVY-FTGNFLDILSARTG-----E 61
 7 LAVETWLLAVTLVLLVLYGTHSHGLFKKLGIPGPTPLPLGN---ILSYRKGFWTFDME 63
 62 BHAKYREKYGSTLRPAGIAGAPVLNSTDPKVNHN-KEAYD-----YPKPGMAARVLR 113
 64 CYKKYKGMWG---YDG--RQVLAITDPIIKTVLKECYSTFNRRRFGVGLKKAL 118
 114 RIATGQGVVTAGEAKHRRIRIMIPSLAQAVKSMVPIFLEKGMELVDMKMDAAEKQVA 173
 119 SI-----AEDEWKIRLSLSPFTSGRLKEMFPII AKYGDVLRNLNREA----- 164
 174 VESAGEKATRIET-EGVDVQKDWGATLDVWALAGFYKSDSLQNTKTNLYVAFVGLT 232
 165 -----ETGKFTVTKDVGAYSMVDVITSTSFQVNDISLNNPD----- 201

```

QY      170 KDMVGSAGEKKATRLTEGVDDVDKDWGRATLDVMALAGFDYKSDLSQNKTNELYVAFV 229
       :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      166 KG-----KPVTMKDVLGAYGMDVITSTSGVNVDSSLNPDPEFEKAK 208
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      230 GLT--DCGFAPTLDSFKRAIMWDFV-PYFRMTKRHRHBIPLUTQGLAVSRRRVGIELMEQKQAV 286
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      209 KLLRFDFDPLL--FSWLFPFLTPVYEML-----NICMPFKDSIEFFKK---F 252
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      287 LGSASDAQVKDVQGQRDLISLLVRA-NTAANLPESOKLSDEEVLAQIISNLIFAGYETSS 345
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      253 VDRMKEKSLDSKQHRVDFQLQMNSHNNSDKDVKSHKALSOMEITAQSIIIFIFAGYETTS 312
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      346 TVLTWMFHRLSDEKAVQDKLRBEICQI--DTDMPTLDDELNALPYLEAFVKSLSRLDPSP 403
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      313 STLSETLHLATHPDIOKKLQDEIDALEANKAPTVDVTWEMEXYLDVNLNETLALYPAN 372
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      404 YANRECLKDDEFPLAEPIVGRDGSVINEVRTTKGTWMLPLFNINRSKFYIGEDAREFR 463
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      373 RLERVCCKDVE-----LNQVIYPRGSTTWIPSYALHHPDQHWSB-PBEFQ 416
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      464 PERWLEDVTDLSNS-IEARYGHQASFIISGRACFGWRFAVAKMFLEVTVRRYVFQFPPII 522
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      417 PERFSKNKGSIDPVYVLPFG-----NGPRNCLGMRFALMMMKLATLKIQNFSPGPC- 469
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
QY      523 SHPEYEHTLIISRPRIVGREK 544
       :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      470 ---KETQIPLKSRQGLLOFEK 489
       :   :   :   :   :   :   :   :   :   :   :   :   :   :

RESULT 11
JX0062
Cytochrome P450 3A7 - human
N:Alternate names: cytochrome P450 3A3; cytochrome P450 HFL33; cytochrome P450 H1p2; cy
N:Contains: oxidoreductase (EC 1.-.-.)
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 28-Jul-2000
C:Accession: JX0062; PX0014; S04983; S02152
R:Komori, M.; Nishio, K.; Ohi, H.; Kitada, M.; Kamataki, T.
J. Biochem. 105, 161-163, 1989
A>Title: Molecular cloning and sequence analysis of cDNA containing the entire coding r
A:Reference number: JX0062; MUID:89255154; PMID:2722762
A:Accession: JX0062
A:Molecule type: mRNA
A:Residues: 1-503 <KOM>
A:Cross-references: GB:D00408; NID:G220148; PIDN:BAA00310.1; PID:G220149
A:Experimental source: fetal liver
R:Komori, M.; Hashizume, T.; Ohi, H.; Miura, T.; Kitada, M.; Nagashima, K.; Kamataki, T.
J. Biochem. 104, 912-916, 1988
A>Title: Cytochrome P-450 in human liver microsomes: high-performance liquid chromatogr.
A:Reference number: PX0012; MUID:89214010; PMID:3243766
A:Accession: FX0014
A:Molecule type: protein
A:Residues: 1-25 <KO3>
A:Experimental source: liver microsome
R:Wrighton, S.A.; Vandenbranden, M.
Arch. Biochem. Biophys. 272, 219-225, 1989
A>Title: Isolation of a new human fetal liver cytochrome P450 cDNA clone: evidence for
A:Reference number: S04983; MUID:89286124; PMID:2786707
A:Accession: S04983
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 282-503 <KO3>
R:Wrighton, S.A.; Vandenbranden, M.
Arch. Biochem. Biophys. 268, 144-151, 1989
A>Title: Isolation and characterization of human fetal liver cytochrome P450H1p2: a thi
A:Reference number: S02152; MUID:89104413; PMID:2492179
A:Accession: S02152
A:Molecule type: protein
A:Residues: 1-11, X', 13-25, X', 27-30 <WRI>
C:Genetics:
A:Gene: GDB:CYP3A7
A:Cross-references: GDB:I34409
A:Map position: 7q22.1-7q22.1

```



```
Db 164 A-----ETGKPVTKKVGAGYSDVITSTSPGVNVDLSLNNPKDPV 204
QY 220 -KTNEL-----YVAFVGLTDFGAPTLDSEKAINW--DFVPYPTWKGRHEIPL 264
Db 205 EKTKLLRFDFDLFSLVFLPELT---PIVEMLNICMFPKDSIEFFK-----250
QY 265 TQGLAVSRVIELMEQKQAVLGASDAQVDKDVQGRDILSLVVRANIAANLPESQ-K 323
Db 251 -----KFYRMKTRLOSVOKRV-----DFLQIMNAHNSDKKESHTA 290
QY 324 LSDEEVLAQISNLLFAGYETSSVLTWMFRLSDEKAVQDKLREICQI--DTDMPTLDE 381
Db 291 LSDMEITAQSIIFIFAGYEPTSTLSFVLHSLATHPTQKKLOEIDRALPNKAPTYDT 350
QY 382 LNALPYLEAFVKESLRLDPPSPYANRCLDKDEDIPLAEPVIGRDSGVINEVITKGTWV 441
Db 351 VMEYELDMVLNETLRLPIGNRLRVCCKDVE-----INGVFWPKGSVV 395
QY 442 MLPLFNINRSKFIYGEDAEERPERWLEDVTDLSNS-IEAPYGHQASFGSPRACFGWRF 500
Db 396 MIPSYALHRDQHWPE-PEERPERFESKENKGSIDPVVLPFG-----NGPRNCIGMRF 448
QY 501 AVAENKAFIVTLRRVQPEP 520
Db 449 ALMNMKALTKVLQNFSPQ 468

RESULT 15
S0892
cytochrome P450 3A16 - mouse
N:Contains: oxidoreductase (SC 1.1.1.1)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 09-May-1997 #text_change 28-Jul-2000
C:Accession: S0892
R:Ittoh, S.; Satch, M.; Abe, Y.; Hashimoto, H.; Yanagimoto, T.; Kamataki, T.
Eur. J. Biochem. 226, 877-882, 1994
A:Title: A novel form of mouse cytochrome P450 3A (Cyp3a-16). Its cDNA cloning and expression
A:Reference number: S0892; MUID:95112853; PMID:7813478
A:Accession: S0892
A:Molecule type: mRNA
A:Residues: 1-504 <ITO>
A:Cross-references: EMBL:D26137; NID:g493670; PIDN:BAA05133.1; PID:g666968
A:Experimental source: fetal liver
A>Note: In the authors' translation residues 9-20 do not match the nucleotide sequence
A>Note: the authors translated the codon TTC for residue 464 as Glu
C:Genetics:
A:Gene: CYP3a-16
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;
F;303-465/Domain: cytochrome P450 homology <P45>
F;469/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 12.9%; Score 368; DB 2; Length 504;
Best Local Similarity 25.8%; Pred. NO. 2e-16;
Matches 132; Conservative 100; Mismatches 171; Indels 108; Gaps 21;

QY 68 EXYGTSL-REFAGIAGAPVLNSTDPVFNHVM-KEAY-----DYPKPGMAARVLRATG 118
Db 66 EXYGTWGLFDG--QIPLVITDPTIKNLVKECFSVFTNRQDFPPVGMISKSISL---120
QY 119 DGWVTAEGEBAHRHRIRIMPSLSAQAVKSMVPIFLKGMELVDMQWDAEAKQAVGESA 178
Db 121 -----AKDEWKRYRALLSPFTTSGNLKEMFPVIEQYG-DILVKYLRQEAEGKQFVA---171
QY 179 GEKKATRLTEGVVDKDWVGRATLDMALAGFDYKDSLQNKTNELYVAVGLTDGFAPT 238
Db 172 -----VKDVLGAYSDVITSTTEGVNIDSLNPNED-----PF 203
QY 239 LDSFKAIM-WDF-----VPYFRTRMKRHEIPLAQGLAVSRVRVGIELMEQKQAVLGSA 290
Db 204 VENAKKVLAFDPDPLSLVALFPFLTPYIYEM--LNICMFPKDSIEFFK----FVDRM 256
```

```
QY 291 SDOAVDKDVGQRDILSLVVRANIAANLPESQ-LSDEEVLAQISNLLFAGYETSSVLT 349
Db 257 TENRLDSKQKHRVDFFYLAMEAVNKSQKDSHKALSEIETIQAQSIIFIFAGYETSSILS 316
QY 350 WMFRLSDEKAVQDKLREICQI--DTDMPTLDELNALPYLEAFVKESLRLDPPSPYANR 407
Db 317 FTVYSLATHPTDIOKKLOEIDEALPNKAPTYDTVMAMEYELDMVLNETLRLPIITNRLQ 376
QY 408 ECLKDEDIFPLAEPVIGRDSGVINEVITKGTWVMLPLFNINRSKFIYGEDAEERPERW 467
Db 377 VCKDVE-----INGYIPKGSVTVIIPSVVLHHDQHWPE-PEEFQPERF 420
QY 468 LEDVTDLSNS-IEAPYGHQASFGSPRACFGWRFVAENKAFIVTLRRVQSEPIISHPE 526
Db 421 SKENKGSIDPVVLPFG-----NGPRNCIGMRFALMNMKALIKVLQNFSPQPC----K 470
QY 527 YEHTTLIISRPRIVGREKEGYQMRLO-VKPV 556
Db 471 ETQIPKLSR-----ELLQPVKPI 490
```

Search completed: April 2, 2004, 14:01:28
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 11 Seconds
(without alignments)
2636.643 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852

Sequence: 1 MFILVLTGALGAFAFWAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414	14.5	503	1	CP3T_PIG
2	413	14.5	503	1	CP39_RAT
3	408.5	14.3	503	1	CP30_SHEEP
4	404	14.2	503	1	CP3C_CAMEL
5	402	14.1	501	1	CP36_RABBIT
6	399	14.0	503	1	CP33_HUMAN
7	398	14.0	503	1	CP3D_MOUSE
8	396.5	13.9	502	1	CP34_HUMAN
9	395.5	13.9	503	1	CP38_MACACA
10	394	13.8	504	1	CP3B_MOUSE
11	391	13.7	504	1	CP31_MOUSE
12	385.5	13.5	503	1	CP37_HUMAN
13	385	13.5	502	1	CP35_HUMAN
14	385	13.5	507	1	CP3S_BOVIN
15	377.5	13.2	503	1	CP3L_CALJA
16	371.5	13.0	501	1	CP3V_MESAU
17	370	13.0	504	1	CP31_RAT
18	368	12.9	504	1	CP3G_MOUSE
19	362	12.7	503	1	CP3F_CAVPO
20	359.5	12.6	511	1	CP4B_MOUSE
21	356.5	12.5	496	1	CP30_FUNHE
22	355.5	12.5	503	1	CP3E_CAVPO
23	351.5	12.3	496	1	CP36_FUNHE
24	351	12.3	520	1	CP58_HUMAN
25	350.5	12.3	1054	1	CYPE_BACSU
26	350	12.3	504	1	CP32_RAT
27	350	12.3	526	1	CP5R_RAT
28	349.5	12.3	524	1	CP41_RAT
29	348.5	12.2	511	1	CP4B_HUMAN
30	346.5	12.1	511	1	CP4B_RAT
31	346.5	12.1	524	1	CPFC_HUMAN
32	345.5	12.1	520	1	YRVA_CAMEL
33	344.5	12.1	502	1	C340_ORILA

34	344	12.1	520	1	CP33_HUMAN	Q08477 homo sapien
35	344	12.1	524	1	CP3B_HUMAN	Q9bb16 homo sapien
36	340.5	11.9	524	1	CP72_CATRO	Q05047 catharanthu
37	340	11.9	503	1	CP3P_MOUSE	Q09158 mus musculu
38	338.5	11.9	503	1	C343_HUMAN	Q9bb55 homo sapien
39	337	11.8	496	1	C4AB_DROME	Q46054 drosophila
40	337	11.8	503	1	CP3H_CAVPO	Q64409 cavia porce
41	337	11.8	520	1	CP32_HUMAN	P78329 homo sapien
42	336.5	11.8	518	1	CP3R_ONCMY	O42563 oncorhynchu
43	336	11.8	522	1	CP34_RAT	P51869 rattus norv
44	334.5	11.7	509	1	YRVA_CAMEL	Q27516 caenorhabdi
45	334.5	11.7	580	1	C973_ARATH	O23365 arabidopsis

ALIGNMENTS

RESULT 1

CP3T_PIG	CP3T_PIG	STANDARD;	PRT;	503 AA.
ID	CP3T_PIG			
AC	P79401;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome P450 3A29 (EC 1.14.14.1) (CYPIIA29)			
GN	CYP3A29.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Small intestine.			
RX	MEDLINE=98347363; PubMed=9682441;			
RA	Nissen P.H., Winteroe A.K., Fredholm M.;			
RT	"Mapping of porcine genes belonging to two different cytochrome P450 subfamilies";			
RL	Anim. Genet. 29:7-11(1998).			
CC	FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.			
CC	CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.			
CC	SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.			
CC	INDUCTION: P450 can be induced to high levels in liver and other tissues by various foreign compounds, including drugs, pesticides, and carcinogens.			
CC	SIMILARITY: Belongs to the cytochrome P450 family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; Z93099; CAB07513.1; ..			
DR	HSP; F14779; IUP2.			
DR	InterPro; IPR001128; Cytochrome_P450.			
DR	InterPro; IPR008072; EP450_CYP3A.			
DR	Pfam; PF00067; P450; 1.			
DR	PRINTS; PR01689; EP450IICYP3A.			
DR	PRINTS; PR00385; P450.			
DR	PROSITE; PS00086; CYTOCHROME_P450; 1.			
KW	Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;			
KW	Microsome; Endoplasmic reticulum.			
FT	METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).			
SQ	SEQUENCE 503 AA; 57198 MW; E3D411B2674FD17F CRC64;			
Query Match	14.5%; Score 414; DB 1; Length 503;			

Best Local Similarity 27.4%; Pred. No. 8.9e-17;
Matches 158; Conservative 85; Mismatches 197; Indels 136; Gaps 23;

QY 12 GLAASFWSAIA--FSLYLAPERS-----SLYLQPNHNTYFTGNELDILSARTGEH-- 63
Db 6 GSTETWULATLSVLVLYGYTHGLFKKIGIPGRLPYF-GN---ILGYKGVDPHD 61
QY 64 AKYREKYGSTLRFAGI--AGAPVLNSTDPKVNHYM-KEAY-----DYPKPGMAARVL 113
Db 62 KXCFQOYG--KMGVYDGRQPLLAATDPNMLKSLVKECYSVFTNRRSFGPLGAMENAL 118
QY 114 RIATGCVVTAAGEAHKRRIMISLSAQAVKSMVPIFLEKGMELVDKXMEDAEDKMA 173
Db 119 SL-----ADDEWKRIITLSLFTFTSGKEMFPIISHYGDLLVSNLAKE-AEG-- 167
QY 174 VGESAGEKKATLETGVGVKDVGRATLDVNALAGPDYKSDSLQNK-----TNELVAF 228
Db 168 -----KPVMTKDI FGAYSMDVITSTAFGWNIDSLNNPDPEVNSKKLLK 212
QY 229 VGLTDGAPATLDSFKALMDF-----VPYF-RTMKRHEIPTOGLAVSRVG 275
Db 213 FSFFDPLSLTFFPPLTFIFEVNITLFPKSSVNFPTSKVRKESBLT----- 262
QY 276 IELMEQOKQAVLGASDAQVDKQVGRDILSLVRANIAANLPSQKLSDEVLQISN 335
Db 263 -----DQOKRV-----DLLQLMINSONSKEMDPKSLSNEELVAQGI 301
QY 336 LIFAGVETSTVLTWPHRLSDKAVODKLRREICQI--DTDMPTLDLNLALPYLEAFVK 393
Db 302 FIFAGVETSSALLAYELATHPDVQKLOBEIBATFNKAPFYDALAQMEVLDVWN 361
QY 394 ESRLDPPSPYANRECLDKEDFIPLAEPVIGRDGVSINEVRITKGTMTWMLPLFNINRSKF 453
Db 362 ETLLPLPIAARLBARACKDVE-----IHGVFPKGTVVVVFVHLRDPD 406
QY 454 IYGEDAEERPERLWEDVDLSNLSIAPYCHQASISGPRACFGWRFAVAKFLFTVL 513
Db 407 LWPE-PEEPERPERFSKKHDTIN---PYTY-LPFGTGRNCGMRFAALMKNLALVRVL 460
QY 514 RYQVQFEPI-----ISHPEYEHITLIISR 536
Db 461 QNFSFKCKETQIPLKLTQGLTQPKPVVLKILPR 496

RESULT 2
CP39 RAT
ID CP39 RAT STANDARD; PRT; 503 AA.
AC P51538; Q64557; Q64631;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome P450 3A9 (EC 1.14.14.1) (CYP11A9) (P450-OLF3) (Olfactive)
DE (3AH15).
GN CYP3A9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97144501; PubMed=8990268;
RA Mainke A., Strockamp D., Roos P.H., Hanstein W.G., Chabot G.G.,
Nef P.;
RT "Expression and inducibility of cytochrome P450 3A9 (CYP3A9) and
other members of the CYP3A subfamily in rat liver.";
RL Arch. Biochem. Biophys. 337:62-68(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96220175; PubMed=8660329;
RA Wang H., Kawashima H., Strobel H.W.;
RT "cDNA cloning of a novel CYP3A from rat brain.";

Biochem. Biophys. Res. Commun. 221:157-162(1996).
CC -!- FUNCTION: This isozyme seems to be implicated in olfaction. Active
in the demethylation of erythromycin as well as benzphetamine.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: Mainly expressed in olfactory epithelium.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U60085; AAB03662.1; -;
CC EMBL; U46118; AAC52582.1; -;
CC PIR; JC4702; JC4702.
CC HSP; P14779; 1JPZ.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008072; E450_CYP3A.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PRO1689; EP450ICYP3A.
CC PRINTS; PRO0385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Olfaction.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 442 457 F -> V (IN REF. 2).
SQ SEQUENCE 503 AA; 57811 MW; 9BBB13E690675EB4 CRC64;
Query Match 14.5%; Score 413; DB 1; Length 503;
Best Local Similarity 27.4%; Pred. No. 1e-16;
Matches 145; Conservative 86; Mismatches 178; Indels 120; Gaps 22;

QY 54 ILSATG-EEIAKY-REXKGTSLFAGI--AGAPVLNSTDPKVNHYM-KEAY----- 101
Db 50 ILAYKGFWEFDKCHKYK--KLVGLYDGRQVLAITDPDIKTIVLVEKCTFTNR 106
QY 102 DYPKPGMAARVLRIATGQVVTABGEAHKRRIMISLSAQAVKSMVPIFLEKGMELVD 161
Db 107 NFGPVGILKALSI-----SEDEWKIRALLSPTFTSGKLKEMFPII----- 149
QY 162 KWMDAAEKDMVAVGESAGEKKATLETGVGVKDVGRATLDVNALAGPDYKSDSLQNK 221
Db 150 NQYTDMLVRNMRQSGSEKGP-----TSMDKIFGAYSMNVITATSGVNVDSLNNPQ 200
QY 222 NELYVAFVGLTDGAPATLDSFKALM-WD-FVPYFRTMKRRHEI-PLTQGLAVSR----- 272
Db 201 D-----PFVEKVKKLLKFDIPLFLSVTLFPFLTFLPEALNVSMFPDVI 246
QY 273 ---RVGIELMEQOKQAVLGASDAQVDKQVGRDILSLVRANIAANLPSQKLSDEEV 329
Db 247 DPFKTSVERMKENR-----MKEKQRMDFLQLMINSONSKYKDSHKAISLVEI 295
QY 330 LAQISNLLPAGYETSTVLTWPHRLSDKAVQDKLRREICQIDTDMP-----TLDELNA 384
Db 296 VAQSVIFPAGYETTSALLSVFLVLLAIHPDIQKLQDE---IDAAPNKAHATYDTLLQ 352
QY 385 LPYLEAFVKESLRLDPPSPYANRECLDKEDFIPLAEPVIGRDGVSINEVRITKGTMTWMLP 444
Db 353 MEYLDVMVNNETLRLYPIAGRLERVKCTDVE-----INGVFIPKGTVMVIP 397
QY 445 LFNINRSKIYGEDAEERPERLWEDVDLSNLSIAPYCHQASISGPRACFGWRFAV 504
Db 398 TFAHKKDPH-YWPEPEERPERFSKKHDTIN---PYTY-LPFGTGRNCGMRFAALMN 451
QY 505 MKAFLEFVTLRRVQFEPI-----ISHPEYEHITLIISRPIV 540
Db 452 MKVALFRVLQNTSFQPKETQIPLKSLQGLLOPKPLLLKWSRDET 500

```

RESULT 3
CP30_SHEEP STANDARD; PRT; 503 AA.
ID Q29496;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A24 (EC 1.14.14.1) (CYP11A24).
GN CYP3A24.
OS Ovis aries (Sheep).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Ching M.S., Chun-Jing J., Ghabrial H., Wooley P.J., Smallwood R.A.,
RA Morgan D.J.;
RT "Ovine foetal liver CYP3A24."
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U59378; AB02657.1; --
DR HSP; F14779; IJFZ.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450ICYP3A.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57360 MW; 804223EAD0304238 CRC64;

Query Match 14.3%; Score 408.5; DB 1; Length 503;
Best Local Similarity 28.5%; Pred No. 1.8e-16;
Matches 165; Conservative 82; Mismatches 197; Indels 135; Gaps 27;

7 LTGALGLAFAFWASIAFFSLYLAPRRS----SLYNLQGNHNYFTGNFLDILSARTG-- 60
3 LIPSPSLETWLLALSLVLLVLYGYSHGLFKLGVSQPRLPYF-GN---VLSYKGVK 58
61 ----FEHAKYKYSTLRERAGIAGAPVNLSTDPKFNHVM-KEAYD-----YKPGM 108
59 EFDEECFKYKQWGVG---FEG--KQPLVITDIPVITKVLVRECYSVFTNRRVFGPMGI 113
109 AARVLRIATGCVVTAEGEAKHRRIRIMPSLSAQVAKSMVPIFLEKGMELVDRMEDAA 168
114 MK-----NAVSVAEDEQWKIRITLLSPFTSGKLKDMFPIIGKYGDVLVRLNLRKE-A 164
169 EKDMVAGSAGEKKATRLTEGTVDVQKDWVGRATLDVWALAGFDYKDSLQNKINELVAF 228

```

```

Db 165 EK-----GKS-----VNMKDFGAVSMDVITSTSGVNIIDSLGNP----- 199
Qy 229 VGLTDFGAPT-----LDSF--KATWDF-VPYER---TMKRREHP-LTQGLAVS 271
Db 200 ---QDFVENAKLRLNFNLDPLLSVFLPFLVFIIEVNLITMPPKSAVDFLTK--SVK 254
Qy 272 RRVGIELMEQKQAVLGSASDAQVDKQVQRDILSLVRANIANLPESOKLSDEEVL 331
Db 255 RIKESRLKDNQPRV-----DFQLMINSONSKETDNHKAISQDELMA 297
Qy 332 QISNLLPAGYTSSTVLTWFMHRLSEDKAVQKLBEEICQIDTDM-----FTLDELNALP 386
Db 298 QSVIFIFAGYETTSNTLSFLLYLATHPDVQKQLOE---IDATPPNKAPTYDVLQOME 354
Qy 387 YLEAPVKESLRLDPPSPYANRECLKDEDFIPLAEPVIGRDSVINEVRITKTMVMLPLF 446
Db 355 YLDMVNETLRMFPIAIVRLDLCKKDE-----IHGVSIPKGTAVTVEIF 399
Qy 447 NINRSKIYGEDAEPRPRWLEDTVDLSNS-IEAPYGHQASFISGPRACGWRFAVEM 505
Db 400 VLHRDPQLWPE-PEEPRPERFSKKNKDSINPYVILPFG-----TGPNCIGMRFAIMNM 452
Qy 506 KAFLEVTLRRVQVEPIIISHPEYEHITLIISRPRIVGREK 544
Db 453 KLAIVRLQNFSPKPC-----KETQIPLKINSQGLIRPEK 487

RESULT 4
CP3C_CANFA STANDARD; PRT; 503 AA.
ID CP3C_CANFA
AC P24463; 1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A12 (EC 1.14.14.1) (CYP11A12) (P450-PBD-1).
GN CYP3A12.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RC MEDLINE=91159488; PubMed=2001406;
RA Ciaccio P.J., Graves P.E., Bourque D.P., Glimsman-Gibson B.,
RA Halpert J.R.;
RT "cDNA and deduced amino acid sequences of a dog liver cytochrome
RT P-450 of the IIIA gene subfamily."
RL Biochim. Biophys. Acta 1088:319-322(1991).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54915; CAA38687.1; --
DR PIR; S14275; S14275.

```

```

DR HSP; P14779; 1JPZ.
DR InterPro; IPR001128; Cytochrome P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP4501ICYP3A.
DR PRINTS; PR00385; P450.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57721 MW; 521710395B5D87 CRC64;

Query Match 14.2%; Score 404; DB 1; Length 503;
Best Local Similarity 29.0%; Pred. No. 3.3e-16;
Matches 146; Conservative 83; Mismatches 184; Indels 90; Gaps 20;

QY 69 KYGSTLRPAGIAGAPVLSNTPDKFVNHVM-KEAYDYPKPGMAARVLIATG-----DGV 121
D 69 KYGSTLRPAGIAGAPVLSNTPDKFVNHVM-KEAYDYPKPGMAARVLIATG-----DGV 121
D 67 KYGRMMGFVD-GRQPVLAITDPMIKTVIVKECY-----SVFNRTTLPVGFPMKSAI 118
QY 122 VTAEGEAKHRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKQAVGESAGEK 181
D 122 VTAEGEAKHRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKQAVGESAGEK 181
D 119 SLSEDEEMKRMRTLLSPFTTCKLKEMFPPIIGQYGDVLVNNLRKE-AEKGA----- 169
QY 182 KATRLTEGVYKVDWVGRATLDVMAALAGFDYKSDSLQNKTNELVAFVGLTDGFAPTLDS 241
D 182 KATRLTEGVYKVDWVGRATLDVMAALAGFDYKSDSLQNKTNELVAFVGLTDGFAPTLDS 241
D 170 -----INLKVFGAYSMVDVITSTSGVNIIDSLNHPQD-----PFVEN 206
QY 242 FKAIM-WDFV-PVFRMTKGRHERIPITQGLA-VSRVVGIELMEQKQAVGLSASQAVDK- 297
D 242 FKAIM-WDFV-PVFRMTKGRHERIPITQGLA-VSRVVGIELMEQKQAVGLSASQAVDK- 297
D 207 TKLLKAFDLDLPPF-----PSILLFPPLTPVFETLNIWLPFKVTDFFRKSVRMKESR 260
QY 298 -KDVQGR--DILSLVRANIANLPSQKSLDEEVLAQISNLLPAGYETSTVLTWVHR 354
D 298 -KDVQGR--DILSLVRANIANLPSQKSLDEEVLAQISNLLPAGYETSTVLTWVHR 354
D 261 LKDKQKRVDFQLMINQSKEMDTHKALSLELVAQSIIFIFAGYETSTVLTWVHR 320
QY 355 LSEDKAVQDKLREICQI--DIDMPTLDLNPALPYLEAFVRESLRDLPPSPYANRECLD 412
D 355 LSEDKAVQDKLREICQI--DIDMPTLDLNPALPYLEAFVRESLRDLPPSPYANRECLD 412
D 321 LATHPDVQKQLEIDATFPNKALPTYDALVQMEYLDVNLNLTLYIAGLERVCKD 380
QY 413 EDIPLAEPVIGRDSGVINEVRITKTMWLPFNINRSKFTYGEDAEFFPERWLEDVT 472
D 413 EDIPLAEPVIGRDSGVINEVRITKTMWLPFNINRSKFTYGEDAEFFPERWLEDVT 472
D 381 VE-----ISGVIPKGVVMTFTLHRDQSLWPE-PSEFPERFSRKNK 424
QY 473 DLSNLSIAPYCHQASPIGPRACFGRFAVAEMKAFVTLRRVQFEPFIISHPYEHITL 532
D 473 DLSNLSIAPYCHQASPIGPRACFGRFAVAEMKAFVTLRRVQFEPFIISHPYEHITL 532
D 425 DSNL-----PTY-LPQTGPRNCIGNRFALMNKALVRLVQLNFSFKPC----KETQIPL 475
QY 533 IISRPVIGREKGYQMRLOVKP 555
D 533 IISRPVIGREKGYQMRLOVKP 555
D 476 KLNAAQIIQPEK---PIVLKVEP 495

RESULT 5
CP36_RABIT STANDARD; PRT; 501 AA.
AC P11707; Q29506;
DT 01-OCT-1989 (Rel. 12, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A6 (EC 1.14.14.1) (CYP11A6) (P450-3C).
GN CYP3A6.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86163152; Pubmed=3349903;
RA Dalet C., Clair P., Dajut M., Fort P., Blanchard J.-M., Maurel P.;
RT "Complete sequence of cytochrome P450 3c cDNA and presence of two
RT mRNA species with 3' untranslated regions of different lengths.";
RL DNA 7:39-46(1988).

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=89380226; Pubmed=2777787;
RA Potenza C.L., Pendurthi U.R., Strom D.K., Tukey R.H., Griffin K.J.,
RA Schwab G.E., Johnson E.F.;
RT "Regulation of the rabbit cytochrome P-450 3c gene. Age-dependent
RT expression and transcriptional activation by rifampicin.";
RL J. Biol. Chem. 264:16222-16228(1989).
CC -!- FUNCTION: Exhibits progesterone 6 beta-hydroxylase activity.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By rifampicin.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M19139; AAA31430.1; -.
CC EMBL; J05034; AAA31178.1; -.
CC PIR; A29487; A29487.
CC PIR; A34236; A34236.
CC HSP; P14779; 1JPZ.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP4501ICYP3A.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 440 440 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 54 55 GI -> VIN (IN REF. 1).
FT CONFLICT 165 169 KIPVD -> ASPST (IN REF. 1).
FT CONFLICT 393 393 MISSING (IN REF. 1).
FT CONFLICT 494 495 RD -> ES (IN REF. 1).
SQ SEQUENCE 501 AA; 57450 MW; 251D21061863ACFB CRC64;

Query Match 14.1%; Score 402; DB 1; Length 501;
Best Local Similarity 27.8%; Pred. No. 4.3e-16;
Matches 156; Conservative 88; Mismatches 200; Indels 118; Gaps 24;

QY 19 ASIAFFSLYLAPRRSSLYNQPNHTNY-FTGNFLDILSARTG--EBHAKYREKYGSTLR 75
D 19 ASIAFFSLYLAPRRSSLYNQPNHTNY-FTGNFLDILSARTG--EBHAKYREKYGSTLR 75
D 15 ASLVLLLYLGYSTHGLFKKMGIPGTPLPFIPTG--ILEVRKGIWDPDIECRKKG--K 68
QY 76 FAGI--AGAPVLNSTDPKVFHVM-KEAY-----DYPKGMARVLRATGDGVVTA 125
D 76 FAGI--AGAPVLNSTDPKVFHVM-KEAY-----DYPKGMARVLRATGDGVVTA 125
D 69 MMGLFDGRQPLMTVDPMIXTLVKECYSVFTNRRSFGVPVGMKAVSI-----SE 120
QY 126 GEAKHRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKQAVGESAGEKATR 185
D 126 GEAKHRRRIMIPSLSAQAVKSMVPIFLEKGMELVDKQMEDAAEKQAVGESAGEKATR 185
D 121 DEDKRVTLISPTFTSGKLKEMLIPIAQGDVLV-KNLRQAEKG----- 165
QY 186 LETEGVDKDWVGRATLDVMAALAGFDYKSDSLQNKTN-----LYVAFVGLTDGAP 237
D 186 LETEGVDKDWVGRATLDVMAALAGFDYKSDSLQNKTN-----LYVAFVGLTDGAP 237
D 166 ---KPYDLKEIFGAYSMVDVITGTSGVNIIDSLNRPQDFVKNVRLKFSF-----FDP 216
QY 238 TLDSFKAIMWDFV-PVFRMTKGRHERIPITQGLAVSRVVGIELMEQKQAVGLSASQAVDK 297
D 238 TLDSFKAIMWDFV-PVFRMTKGRHERIPITQGLAVSRVVGIELMEQKQAVGLSASQAVDK 297
D 217 LLLSI-----TLFPFL-----TFEALHIS-MPKQVMDFLKTSVEKIDRLKDK 262
QY 298 KDVQGRDILSLVRANIANLPSQKSLDEEVLAQISNLLPAGYETSTVLTWVHR 357
D 298 KDVQGRDILSLVRANIANLPSQKSLDEEVLAQISNLLPAGYETSTVLTWVHR 357
D 263 QK-RRVDFQLMINQSKEMDTHKALDDIEVAQSIILFAGYETSTVLTWVHR 321
QY 358 DRAVQDKLREICQIDTMP-----TLDELNALPYLEAFVRESLRDLPPSPYANRECLD 412
D 358 DRAVQDKLREICQIDTMP-----TLDELNALPYLEAFVRESLRDLPPSPYANRECLD 412

```


Db 322 HDVQOKQER---IDTLLPKKELATYDTLVKMEYLDMMVNETLRLYPIAGRLERVCCKD 378
 Qy 413 EDFIPLAEPIVGRDGSVINEVRIKGTMTVMPLFNINRSFIYCEDAEAFRRERMLEDTV 472
 Db 379 VD-----INGTIFPKGTIVVMPTVALHRDPQHWTPE-PDFRFRERFSKKNK 422
 Qy 473 DLSNS-IEAPYGHQASFTSGPRACFGRRPAPVAEMKAFVTLRVOFE-----PI--- 521
 Db 423 DMINPIYVHPFG-----AGPRNCLGMRFALMNKIALVRLMONTFGRFKCKETQVPLKLG 476
 Qy 522 ---ISHPEYEHITLIISRPRIV 540
 Db 477 KQGLLQPEKPIVLKVVSRDGI 498

RESULT 6
 CP33_HUMAN STANDARD; PRT; 503 AA.
 AC P05184;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 3A3 (EC 1.14.14.1) (CYP3A3) (HLP).
 GN CYP3A3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=86259780; PubMed=3460094;
 RA Molowa D.T., Schuetz E.G., Wrighton S.A., Watkins P.B., Kremers P.,
 RA Mendez-Picon G., Parker G.A., Guzelian P.S.;
 RT "Complete cDNA sequence of a cytochrome P-450 inducible by
 RT glucocorticoids in human liver.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5311-5315(1986).
 RN [2]
 RP SEQUENCE OF 1-20.
 RC TISSUE=Liver;
 RX MEDLINE=85298342; PubMed=3898085;
 RA Watkins P.B., Wrighton S.A., Maurel P., Schuetz E.G., Mendez-Picon G.,
 RA Parker G.A., Guzelian P.S.;
 RT "Identification of an inducible form of cytochrome P-450 in human
 RT liver.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6310-6314(1985).
 CC -I- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -I- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -I- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -I- INDUCTION: By glucocorticoids.
 CC -I- SIMILARITY: Belongs to the cytochrome P450 family.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; D00003; BAA00001.1; -.
 EMBL; M13785; AAA35742.1; -.
 PIR; A29410; A29410.
 HSSP; P14779; 1JPZ.
 DR Genew; HGNC:2636; CYP3A3.
 DR GO; GO:0005792; C-microsome; TAS.
 DR GO; GO:0015034; F-cytochrome P450 activity; TAS.
 DR InterPro; IPR001128; Cytochrome_P450.

DR InterPro; IPR008072; EP450_CYP3A.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01689; EP450IICYP3A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT INIT_MET 0 0
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 503 AA; 57428 MW; 9B85DIF729658FC0 CRC64;

Query Match 14.0%; Score 399; DB 1; Length 503;
 Best Local Similarity 28.4%; Pred. No. 6.4e-16;
 Matches 156; Conservative 83; Mismatches 182; Indels 128; Gaps 24;

Qy 13 LAAPSW--ASIAFFSLYLAPRRS--SLYNLQ--PHNTNY--FTGNFLDILSARTG-----E 61
 Db 6 LAMETWLLAVSLVLLYLYGTHSHGLFKKIGIPGPTPLPFLGN---ILSYHKFCMFDMDE 62
 Qy 62 EHAKYREKYGSTLRPAGIAGAPVLNSTDPKVFNVHM--KEAYD-----YPKFGMAARVL 113
 Db 63 CHKKYGVWGF---YDG--QQPVLAITDPDMIKLVVKECVSVFTNREPFGVGMKSAI 117
 Qy 114 RATGQGVVTAEGEBAHKRRIRIMIPSLSAQAQVKSMDPIFEKGMELVDOMEDAAEKDWA 173
 Db 118 SI-----ADEEKKRLRSLLSPTFTSGKLEKEMVPIIAQYGDVLRNLRRE----- 162
 Qy 174 VQESAGEKKATRETEGVVDKDWVGRATLDVWALAGPDYKSDSLQNKTNELYVAVGLTD 233
 Db 163 -----RETGKPVTLKDVFGAYSMVDVITSSFGVNVDSLNPPQD----- 200
 Qy 234 GPAPTLDSEKATM-WDFV-PYFRIMK--RRHEIPLTQGL-----AVSRVVG 275
 Db 201 ---PLVETKLLRDFLDLPFLSLTFVFPFLIPLEVINICVFPREVTNLRKAVRMKE 257
 Qy 276 IELMEQKQAVLGSASDAQVDKQVQGRDILSLVRANIAANLPESOK--LSDEEVLAQIS 334
 Db 258 SRLEDTQKRV-----DFLQIMDSHKSKKTESHKALSDELVAQSI 300
 Qy 335 NILFAGYETSSVLTWMPHRLSEDKAVQDKLRREICQI---DTDMPTLDELNALPYLEAFV 392
 Db 301 IFIFAGYETSSVLSFIMYELATHPDVOQKQLEIDAVLPNKAPPTVDTVLQMEYLDMMV 360
 Qy 393 KESRLRDPSPYANRECKLDEDFIAPFPVIGRDSVINEVRIKGTMTVMPLFNINRSK 452
 Db 361 NETLRLFPJALMLERLVCKDVE-----INGMFIPKGVVWVMPYSVALHRDP 405
 Qy 453 FYGEDAEAEFRERLEEDVTSLSNS-IEAPYGHQASFTSGPRACFGRRPAPVAEMKAFV 511
 Db 406 -KYWTEPEKFLPERFSKKNKONIDPIYITPFG-----SGPRNCIGMRFPALMMKLALIR 458
 Qy 512 TLRRVQFEP 520
 Db 459 VLQNFSEKP 467

RESULT 7
 CP3D_MOUSE STANDARD; PRT; 503 AA.
 ID CP3D_MOUSE
 AC Q64464;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 3A13 (EC 1.14.14.1) (CYP3A13).
 GN CYP3A13 OR CYP3A-13.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=dy; TISSUE=Liver;
 RX MEDLINE=95101705; PubMed=7803471;

RA Yanagimoto T., Itoh S., Sawada M., Hashimoto H., Kamataki T.;
RT "Molecular cloning and functional expression of a mouse cytochrome P-
RT 450 (Cyp3a-13): examination of Cyp3a-13 enzyme to activate aflatoxin
RT B1 (AFB1).";
RL Biochim. Biophys. Acta 1201:405-410(1994).
CC -!- FUNCTION: Can activate aflatoxin B1 to a genotoxic product.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O;
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC ENBL; X63023; CAA44754.1; -;
CC PR; S50211; S50211.
CC HSP; P14779; LUP2.
CC MGD; MGI:88610; Cyp3a13.
CC InterPro: IPR001128; Cytochrome P450.
CC InterPro: IPR008072; EP450_CYP3A.
CC Pfam: PF00067; P450; 1.
CC PRINTS; PR01689; EP4501ICYP3A.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57492 MW; 15DDE2606B337FCF CRC64;

Query Match 14.0%; Score 398; DB 1; Length 503;
Best Local Similarity 27.0%; Pred. No. 7.3e-16;
Matches 142; Conservative 74; Mismatches 179; Indels 130; Gaps 20;
QY 63 HAKYREKYSTLRFGAGAPVLNSTDPKVFNVHM-KEAYD-----YKPGMAARVLR 114
DB 65 HKYKXWNG--LYDG--RQPLAITDPDIITKVLKCYSTFTNRRRFGPVGILKXAS 119
QY 115 IATGDGVVTRAGBAHRRHRIIPSLSAQVKSVMPIFEKGMELVDQMDAAEKDQAV 174
DB 120 I-----SENBEWKRIALLSPFTSGLKEMFPFI-----NQFTDVLVRNMQ 162
QY 175 GESAGEKKATRLTEGVKVDKWDVGRATLDVMAAGFDYKSDSLQNT----- 221
DB 163 G--LGECKPT-----SMKIDFGAYSDVITATSGVNIIDSLNNPQDPFVEKIKLKF 213
QY 222 ---NELYAVVGLTDFAPTLDSFKALM--DFVYFRMKRHRIPITQGLAVSRVGI 276
DB 214 DIFDPLFLS-VTLFPFLTPVFDALNSLPFRDVISFPTT-----SVERKEN 259
QY 277 ELMECKKQAVLGASDAQVDKDVQGRDILSLVRANTANLPESQKLSDEEVLQISNL 336
DB 260 RMKEKEKQV-----DFQLQMINSONYKTKESHKALSDVEIVQSVIF 302
QY 337 LFAGYETSSVLTWTFHRLUSEDKAVQDKUREEICQIDTDP-----TLDELNALPYLEAF 391
DB 303 IFAGYETSSALSFALYLLAIHPDVKQLQDE---IDAALPNKAPATYDTLQMEYLDV 359
QY 392 VKESLRDPPSPYANRECKLDEDFIPLAEPVIGRDSVINEVRITKGTWMLPLFNINRS 451
DB 360 VNETRLRYPAGRLERVCKTDV-----INGLFIKGTVMVITPFLHKD 404
QY 452 KFIYGDAEERFRPERMLEVDTSLSNIEAPYHQASFSIGPRACFGWRFAVAKFLV 511
DB 405 P-KYWPPEFRPERFSKQNDGSI---PYMY-LPFGSGPRNCIGMRPALINMKVALVR 458

QY 512 TLRRVQFEPI-----ISHPEYEHITLIISRRIVGRE 543
DB 459 VLQFTVQPCKEVEIPLKSKQGLQFENPPLLKVVSRDETWSDE 503
RESULT 8
ID CP34 HUMAN STANDARD; PRT; 502 AA.
AC P06584; Q16757; Q9UK50;
DT 01-JAN-1988 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A4 (EC 1.14.13.67) (Quinine 3-monooxygenase)
DE (CYP11A4) (Nifedipine oxidase) (NF-25) (P450-PCN1).
DE CYP3A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88195781; PubMed=3267210;
RA Gonzalez F.J., Schmid B.J., Umeno M., McBride O.W., Hardwick J.P.,
RA Meyer U.A., Gelboin H.V., Idle J.R.;
RT "Human P450PCN1: sequence, chromosome localization, and direct
RT evidence through cDNA expression that P450PCN1 is nifedipine
RT oxidase.";
RL DNA 7:79-86(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87041402; PubMed=3464943;
RA Beaune P.H., Umbenhauer D.R., Bork R.W., Lloyd R.S., Guengerich F.P.;
RT "Isolation and sequence determination of a cDNA clone related to
RT human cytochrome P-450 nifedipine oxidase.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:8064-8068(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89108438; PubMed=2563251;
RA Spurr N.K., Gough A.C., Stevenson K., Wolf C.R.;
RT "The human cytochrome P450 CYP3 locus: assignment to chromosome 7q22-
RT qter.";
RL Hum. Genet. 81:171-174(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89093163; PubMed=2463251;
RA Bork R.W., Muto T., Beaune P.H., Srivastava P.K., Lloyd R.S.,
RA Guengerich F.P.;
RT "Characterization of mRNA species related to human liver cytochrome P-
RT 450 nifedipine oxidase and the regulation of catalytic activity.";
RL J. Biol. Chem. 264:910-919(1989).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=99256082; PubMed=10322772;
RA Chen Q., Wu J., Yu Y.;
RT "Establishment of transgenic cell line CHL-3A4 and its metabolic
RT activation.";
RL Zhonghua Yu Fang Yi Xue Za Zhi 32:281-284(1998).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=21163842; PubMed=11266076;
RA Gellner K., Eiseit R., Huestert E., Arnold H., Koch I., Habert M.,
RA Beglmann C.J., Burk O., Bunterfuss D., Becher S., Bishop C.,
RA Koebe H.-G., Brinkmann U., Klenk H.-P., Klein K., Meyer U.A.,
RA Wojnowski L.;
RT "Genomic organization of the human CYP3A locus: identification of a
RT new, inducible CYP3A gene.";
RL Pharmacogenetics 11:111-121(2001).
RN [7]
RP SEQUENCE FROM N.A., AND VARIANTS VAL-117 AND ARG-217.
RX MEDLINE=21098888; PubMed=11181494;
RA Haieh K.-P., Lin Y.-Y., Cheng C.-L., Lai M.-L., Lin M.-S.,

RA Siest J.-P., Huang J.-D.;
 RT "Novel mutations of CYP3A4 in Chinese."
 RL Drug Metab. Dispos. 29:268-273 (2001).
 RN [8]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RA Zhuge J., Qian Y., Xie H., Yu Y.;
 RT "Sequence of a new human cytochrome P450-3A4 cDNA."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP CHARACTERIZATION.
 RA Zhang H., Coville P.F., Walker R.J., Miners J.O., Birkett D.J.,
 RA Wawoncluk S.;
 RT "Evidence for involvement of human CYP3A in the 3-hydroxylation of
 RT quinine."
 RL Br. J. Clin. Pharmacol. 43:245-252 (1997).
 RN [10]
 RP CHARACTERIZATION.
 RA Zhao X.J., Kawashiro T., Ishizaki T.;
 RT "Mutual inhibition between quinine and etoposide by human liver
 RT microsomes. Evidence for cytochrome P4503A4 involvement in their
 RT major metabolic pathways."
 RL Drug Metab. Dispos. 26:188-191 (1998).
 RN [11]
 RP VARIANTS PRO-221 AND THR-444.
 RA MEDLINE=20132287; PubMed=9456308;
 RA Sato F., Sapone A., Elizondo G., Stocker P., Miller V.P., Zheng W.,
 RA Ranaio H., Crespi C.L., Gonzalez F.J.;
 RT "CYP3A4 allelic variants with amino acid substitutions in exons 7 and
 RT 12: evidence for an allelic variant with altered catalytic activity."
 RL Clin. Pharmacol. Ther. 67:48-56 (2000).
 RN [12]
 RP VARIANTS GLN-161; SER-188; PRO-292; THR-444 AND SER-466, AND
 RP CHARACTERIZATION OF THE VARIANTS.
 RX MEDLINE=21571796; PubMed=11714865;
 RA Dai D., Tang J., Rose R., Hodgson E., Bientstock R.J.,
 RA Mohrenweiser H.W., Goldstein J.A.;
 RT "Identification of variants of CYP3A4 and characterization of their
 RT abilities to metabolize testosterone and chlorpyrifos."
 RL J. Pharmacol. Exp. Ther. 299:825-831 (2001).
 RN [13]
 RP VARIANTS ASP-55; GLN-129; ILE-169; HIS-173; MET-362; PHE-372 AND
 RP LEU-415.
 RX MEDLINE=21364010; PubMed=11470997;
 RA Eiselt R., Domanski T.L., Zibat A., Mueller R., Presecan-Siedel E.,
 RA Huster E., Zanger U.M., Brockmoller J., Klenk H.-P., Meyer U.A.,
 RA Khan K.K., He Y.-A., Halpert J.R., Wojnowski L.;
 RT "Identification and functional characterization of eight CYP3A4
 RT protein variants."
 RL Pharmacogenetics 11:447-458 (2001).
 RN [14]
 RP VARIANTS PRO-14; GLN-161; HIS-173; SER-184 AND PHE-372.
 RX MEDLINE=21864884; PubMed=11875366;
 RA Lamba J.K., Lin Y.S., Thummel K., Daly A., Watkins P.B., Strom S.,
 RA Zhang J., Schuetz E.G.;
 RT "Common allelic variants of cytochrome P4503A4 and their prevalence in
 RT different populations."
 RL Pharmacogenetics 12:121-132 (2002).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It performs a variety
 CC of oxidation reactions (e.g. caffeine 8-oxidation, omeprazole
 CC sulfoxidation, midazolam 1-hydroxylation and midazolam 4-
 CC hydroxylation) of structurally unrelated compounds, including
 CC steroids, fatty acids, and xenobiotics. The enzyme also
 CC hydroxylates etoposide.
 CC -!- CATALYTIC ACTIVITY: Quinine + NADPH + O(2) = 3-hydroxyquinine +
 CC NADP(+) + H(2O).
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- TISSUE SPECIFICITY: Expressed in prostate and liver.
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other

CC tissues by various foreign compounds, including drugs, pesticides,
 CC and carcinogens.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
 CC NOTE=CYP3A4 alleles;
 CC WWW="http://www.imm.ki.se/cypalleles/cyp3a4.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M18907; AAA35745.1; -;
 CC EMBL; M14096; AAA35744.1; -;
 CC EMBL; X12387; CAA30944.1; -;
 CC EMBL; J04449; AAA35747.1; -;
 CC EMBL; AF182273; AAF13598.1; -;
 CC EMBL; AF280107; AAG32290.1; -;
 CC EMBL; AF209389; AAF21034.1; -;
 CC PIR; A29815; A29815.
 CC HSSP; P14779; 10PZ.
 CC Genew; HGNC:2637; CYP3A4.
 CC MIN; 124010; -;
 CC GO; GO:0015034; F:cytochrome P450 activity; TAS.
 CC GO; GO:0004497; F:monooxygenase activity; TAS.
 CC GO; GO:0006629; P:lipid metabolism; TAS.
 CC InterPro; IPR001128; Cytochrome P450.
 CC InterPro; IPR008072; EP450_CYP3A.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR01689; EP450IICYP3A.
 CC PRINTS; PR03385; P450.
 CC PROSITE; PS00086; CYTOCHROME P450; 1.
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 CC Microsome; NADP; Endoplasmic reticulum; Polymorphism.
 CC INIT MET 0 BY SIMILARITY
 CC METAL 441 441 IRON (HEME AXIAL LIGAND) (BY SIMILARITY)
 CC VARIANT 14 14 L-> P (in allele CYP3A4*14).
 CC VARIANT 55 55 FTId=VAR_011597.
 CC VARIANT 117 117 FTId=VAR_011598.
 CC VARIANT 129 129 FTId=VAR_011599.
 CC VARIANT 161 161 FTId=VAR_011600.
 CC VARIANT 169 169 FTId=VAR_011601.
 CC VARIANT 173 173 FTId=VAR_011602.
 CC VARIANT 184 184 FTId=VAR_011603.
 CC VARIANT 188 188 FTId=VAR_011604.
 CC VARIANT 217 217 F-> S (in allele CYP3A4*17; exhibits
 CC lower turnover numbers for testosterone
 CC and chlorpyrifos).
 CC VARIANT 217 217 FTId=VAR_014322.
 CC VARIANT 217 217 P-> R (in allele CYP3A4*5).
 CC FTId=VAR_011605.
 CC
 CC Query Match 13.9%; Score 396.5; DB 1; Length 502;
 CC Best Local Similarity 28.7%; Pred. No. 8.9e-16;
 CC Matches 153; Conservative 88; Mismatches 195; Indels 97; Gaps 24;
 CC
 CC QY 13 LAAPSW--ASTAFSLVLAAPRS-SLYNLOG-PNHINY-PTGNFLDILSARTG-----E 61
 CC 6 LAMETVLLVLLVLLVLYGTHSGHGFKKLGIQGPPLPLPLGN---ILSTHGKFCMFDE 62
 CC Db
 CC QY 62 EHAKYKYGSTLRFAGIAGAPVLNSTDPKVFNVHM-KEAYDY---PKFGMAARVLRIAT 117
 CC 63 CHKYKVGWGF---YDG--QQPVLAITDPDMIKTLVKCYSVFTNRRPFGPGFMKSA- 116
 CC Db

7 LTGALGLAASFASIASIAFFSLY-LAPRRSSLYNLOG-----PNNNTYPTGNFLDLSAR 58
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
 3 LVSAALSLETVALLAISLVLLRYGYRKKHLPKQGIPOGPKLDP-----FLGTVLVNYEKL 57
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 59 TGBEHAKYRKYGSTL-REAGTAGAVLNSTDPKVFNVHM-KEAY-----DYKPQGA 109
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 58 WKFDMECYK-KYKGTWGLFDG--QTPELLAVTPETIKNVLKVECSVTNRDRDFGVGIM 114
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 110 ARVRIATGDVVTAEGEAHKHRIRIMPISLSAQAVKSVMPIFLEKGMELVDQMMDAAE 169
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 115 SKAISISKDD-----EWKRYEALLSPFTSGKLKEMFPVIEQVG-DILVKYLROKAK 165
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 170 KMAVGSAGEKKATRLTEFGVDVKDWVGRATLDVMALAGEDYKSDSLONKTNELVAFV 229
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 166 KG-----KPVTKDVGLGAYSMDVITSTSGVNVDLSLNNPEDPPFEKAK 208
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 230 GLT---DGFAPTLSFKAIMDFV-PYRTWKRRHEIPLTQGLAVSRRVGIELMEOKQAV 286
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 209 KLRFDFDFDLL-FSVLPFPLTPVVEML-----NICMFPKDSIEFFKK---F 252
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 287 LGSASDAOVKQVQGRDILSLAVRA-NTAANLPESCKLSDEVLQAISNLLFAGYETSS 345
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 253 VDRWKSERLDSKQHRVDFQLMMNHNSNKVKSHKALSDEITRAQSIIFFAGYEITS 312
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 346 TVLTWMFHRLSEDKAODKLRBEICQI--DTDMPTIDELNALPYLEAFVKESIRLDPPSP 403
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 313 STLSTLHLATHPDIOKKLODSIDEALPNKAPPTYDTVMEMEYLDVMVLNETRLYPFIAN 372
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 404 YANRECUKDEDPIELAEPVIGROSSVINEVRITKGTMVWLPLFNINRSKFIVGEDAEPR 463
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 373 FLERVCKKDVE-----LNGVIIPKGTVMIPSVALHHDPQHSE-PEEFQ 416
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 464 PERMLEDVDTSINS-IBAPYGHQASTISGPRACGWRFAVAENKPAFLFTLARVVOFEPII 522
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 417 PERSKENKSGIDPYVYLPFG-----NGPRNLGMRFAFMNMKLALTIMQNFSQPCC- 469
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 523 SHPEVEHTLIISRPTVGREK 544
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 470 ---KETQPLKLSRQGGLQEK 488
 :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::

RESULT 11
 C341_MOUSE STANDARD; PRT; 504 AA.
 ID_C341_MOUSE
 AC QJMA7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 3A41 (EC 1.14.14.1).
 GN CYP3A41.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN [1]
 RX SEQUENCE FROM N.A.
 RP STRAIN=dqy; TISSUE=liver;
 RC MEDLINE=20239668; PubMed=10775455;
 RA Sakuma T., Takai M., Endo Y., Kuroiwa M., Ohara A., Jarukanjorn K.,
 RA Honna R., Nemoto N.;
 RT "A novel female-specific member of the CYP3A gene subfamily in the
 RT mouse liver.";
 RL Arch. Biochem. Biophys. 377:153-162(2000).
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
 CC (Potential).
 CC -!- TISSUE SPECIFICITY: Expressed in liver. Also expressed in the
 CC kidneys of female mice, with traces in the stomach, ovary, and
 CC heart of female mice and in the testis of male mice.
 CC -!- DEVELOPMENTAL STAGE: Detected immediately after birth in the
 CC livers of animals of both sexes, but increased with age in
 CC females, whereas it was gradually reduced in males, relluting in

CC predominantly female-specific expression in livers.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB033414; BAA95951.1; --
DR MGD; MGI:1858451; CYP3A41.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1
DR PRINTS; PR01689; EP450IICYP3A.
DR PROSITE; PS00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 504 AA; 57959 MW; 2BC645B6E9CD48A CRC64;

Query Match 13.7%; Score 391; DB 1; Length 504;
Best Local Similarity 26.2%; Pred. No. 1.8e-15;
Matches 148; Conservative 100; Mismatches 212; Indels 104; Gaps 23;

QY 7 LTGALGAAFWASIAFFSLY-LAPRRSLYNLQ-PNHT-----NYFTGNFLDIL 55
DB 3 LFSALSDFTWALLAILVLLRYGRTGFLFKQIQIPGTPPLFLGTVLNLYKGLW---- 58

QY 56 SARTGEHAKYREKYGSTL-RFAGIAGAPVLNSTDPKFNHVM-KEAY-----DYKXP 106
DB 59 -----KFDMECYEKYKGTWGLFDG-QMPLFITDPEMIKNVLKCEKSVFTNRRFGPV 111

QY 107 GMAARVLRATGCVTVTAGEAKHRRIMISLSAQAVKSWPIFLKGMELDKWMD 166
DB 112 GIMSKAISL-----SKDEWKRYALLSFTTSGKLEKMPVIEYQDILVKYLMQE 163

QY 167 AAEKDMVAGESAGEKKATRLTEGVVDKDWVGRATLDVNALAGFDYKSDSLQNKTNELV 226
DB 164 -AEKG-----KPVTKMDVLGAYSIDVITSTSGVNVDSLNNPDEPFVE 205

QY 227 AFVGL--TGFAPTLDSFKAIMWDFV-PFRMKRHRHPIPTQGLAVSRVGIELMQKK 283
DB 206 KAKGILRVDFDFDLV--FSVLLFPFLTPVYEML-----NICMPFKDSIEFFKK-- 251

QY 284 QAVLGASDAQVDKQVQGRDILSLVRANIANLPESOK-LSDEVLQAISNLLFAGYE 342
DB 252 --FVNRKESRLDSKQKRVDFQLQMMNHNNSKDKSHKALSDMEITAQSIIVIFAGYE 309

QY 343 TSSTVLTWPHRLSEDAKQVDKLEBICQI--DTDMPTLDELNALPYLEAFVKSLRLDP 400
DB 310 TTSSTLSFTLYCIATHPDIOKQLQEBIDETLKNKAPPTVDTVMEMEYLDWNLNLTSLRP 369

QY 401 PSPVANRECKLDDFTPLAEPVIGROGVSINEVRITKGTVMPLFNINRSKFYIGEDAE 460
DB 370 IGRLERFCKQVE-----LNGVYIPKSTVMIPSYALHHPQHWPE-PE 413

QY 461 EFRPERKLEVDTSLSNIEAPYGHQASFTSGRACFGWRFAVAENKALFVTLRRVQEP 520
DB 414 EFQPERFSKENKGSID----PYLYM-PFGIGPRNCIGNRFAPMTYKLTAKYMNQFSQP 468

QY 521 IISHPEVEHITLIISPRIVGREK 544
DB 469 C-----QETQIPLKLSROGLLQPEK 488

RESULT 12
CP37 HUMAN
ID CP37 HUMAN
AC P24462;

DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A7 (EC 1.14.14.1) (CYP3A7) (P450-HFLA).
GN CYP3A7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1] J. Biochem. 105:161-163(1989).
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=89255154; PubMed=2722762;
RA Komori M., Nishio K., Ohi H., Kitada M., Kamataki T.;
RT "Molecular cloning and sequence analysis of cDNA containing the
RT entire coding region for human fetal liver cytochrome P-450.";
RL J. Biochem. 105:161-163(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21163842; PubMed=11266076;
RA Gellner K., Bisset R., Huestert E., Arnold H., Koch I., Haberl M.,
RA Deglmann C.J., Burk O., Buntefuss D., Escher S., Bishop C.,
RA Koebel H.-G., Brinkmann U., Klein H.-P., Klein K., Meyer U.A.,
RA Wojnowski L.;
RT "Genomic organization of the human CYP3A locus: identification of a
RT new, inducible CYP3A gene.";
RL Pharmacogenetics 11:111-121(2001).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -1- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP3A7 alleles;
CC WWW="http://www.imm.ki.se/CYPalleles/cyp3a7.htm".
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00408; BAA00310.1; --
DR EMBL; AF280107; AAG32289.1; --
DR FTR; JX0062; JX0062.
DR R5SP; P47779; IJFZ.
DR Genew; HGNC:2640; CYP3A7.
DR MIM; 605340; --
DR GO; GO:0015034; F:cytochrome P450 activity; TAS.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1
DR PRINTS; PR01689; EP450IICYP3A.
DR PROSITE; PS00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57470 MW; 087CCEED9BAC314C CRC64;

Query Match 13.5%; Score 385.5; DB 1; Length 503;
Best Local Similarity 29.0%; Pred. No. 3.8e-15;
Matches 157; Conservative 79; Mismatches 191; Indels 115; Gaps 26;

DR MTM; 605325; --
 DR GO:0005792; C:mitochondrion; TAS.
 DR GO:0015034; F:cytochrome P450 activity; TAS.
 DR GO:0004497; F:monooxygenase activity; TAS.
 DR GO:0008202; P:steroid metabolism; TAS.
 DR InterPro: IPR001128; Cytochrome P450.
 DR InterPro: IPR008072; EP450_CYP3A.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Polymorphism.
 FT METAL 441 441 IRON (HEME AXIAL LIGAND)
 FT VARIANT 398 398 /TIGEAR 008385.
 FT SEQUENCE 502 AA; 57108 MW; D5A2302E2638717 CRC64;
 Query Match 13.5%; Score 385; DB 1; Length 502;
 Best Local Similarity 28.1%; Pred. No. 4.1e-15;
 Matches 153; Conservative 94; Mismatches 196; Indels 106; Gaps 25;
 QY 13 LAAPSW-ASIAFSLYL-APRRSSLYNQ-PNHTNY-FTGNFLDILSARTG-----E 61
 Db 7 LAVETWLLAVSLVLLYXGTRTHGLFKRLGIPGTPPLLGN---VLSYRQGLWKFDTE 63
 QY 62 EHAKYREKYSTLRERFAGIAGAVLNSTDPKFNHVM-KEADYDPKPGMAARVLRATG-- 118
 Db 64 CYKYGKMGWT---YEG-QLPVLATDPDVTITLVKCY-----SVFTNRSILGPV 111
 QY 119 ----DGVTAGBAKRRHRRIMTPSLSAQAQVSMVPIFEKGMELVDKMDAEKQNAV 174
 Db 112 GFMKSAISLAEDBEWKRIISLPTFTSGKLKEMFPITIAQYGDVLVRLRRE-AEKG--- 167
 QY 175 GESAGEKATLETEGVGVKDWVGRATLDVMAAGDYKSDSLQ-----KINELYVAF 228
 Db 168 -----KPTVLKIFGAVMDVITGTSFGVNDISNPNQDPFVESIKKFKF 213
 QY 229 VGLTDGAPATLDSFKAIMDVFVYFRMKRHEIPLTQGLAVS---RRVGIELMEQKQAV 286
 Db 214 ----GFLDEL-FLSIILPFL-----TPVFEALNVSLPDKTINFLSKS----- 252
 QY 287 LGSASDAQVKKVQGVQEDILSLAVRANIANLPESQKLSDEVLQAISNLLFAGVETST 346
 Db 253 VNRKKSLRNDKQKRLDFLQMLDSONSKETESHKALSDLELAQSIIFIFAGVETTS 312
 QY 347 VLTWFMHRLSDKAVQDKLEECICQI--DMDPTLDELNALPYLEAFVKESLRDPPSPY 404
 Db 313 VLSFTLYELATHPDVQQLQKEIDAVLPNKAPPTYDAVVQMEYLDVWVNETLRLFPVAIR 372
 QY 405 ANRECLKDEDFPIAEVPIGRDGSVINEVITKGTWMLPLFNINRSKEFIYGEDAEERFP 464
 Db 373 LERTCKKDE-----INGVPIKGMVVIPTVALHDDP-KYWTEPEERFP 416
 QY 465 ERWLEDVTDLSNS-IEAPYGHQASFTSGPRACFGMEFAVAEMKAFILFVTLRRVQEP 520
 Db 417 ERFSKK-KSDIDPIYTFPG-----TGPENCIGRFALNMKALIRVLQNFSEK 466

RESULT 14
 ID CP3S BOVIN STANDARD; PRT; 507 AA.
 AC P79102;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 3A28 (EC 1.14.14.1) (CYP11A28).
 GN CYP3A28.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;

RP SEQUENCE FROM N.A.
 RA STRAIN:Hereford; TISSUE=Liver;
 RA Natsuhoori M., van Raak M., Ligteneberg M., Kleij L., ten Berge D.,
 RA Zweers-Zeilmaker W.M., de Groene E.M., van Miert A.S.J.P.A.,
 RA Witkamp R.F., Horbach G.J.M.J.;
 RT "Isolation of a bovine full length cytochrome P450 (CYP3A) cDNA
 RT sequence and its functional expression in V79 cells.";
 RL Environ. Toxicol. Pharmacol. 3:17-24(1997).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other
 CC tissues by various foreign compounds, including drugs, pesticides,
 CC and carcinogens.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y10214; CAAT1266.1; --
 DR HSP; P14779; IJFZ.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008072; EP450_CYP3A.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT SEQUENCE 507 AA; 58152 MW; FDFD846SCC96DF57 CRC64;
 Query Match 13.5%; Score 385; DB 1; Length 507;
 Best Local Similarity 25.5%; Pred. No. 4.1e-15;
 Matches 140; Conservative 94; Mismatches 192; Indels 122; Gaps 22;
 QY 38 LOGPNHTNYFTGNFLDILSARTGEEHAKYR-----EKYGSTLRERFAGIAGAVLN 87
 Db 38 IFGPRPVYF-----GSTMAVHKGIPEDFNQCFKYGKMGVFE-GRQPMIAI 84
 QY 88 TDPKFNHVM-KEADYD-----YKPGMAARVLRATGCVVTAEGEAHRRRIMPS 139
 Db 85 TPDPIIKTVLVEKCVSVTNRIFGPMIGMKYALSLAW-----DEQWKRIITLSPA 136
 QY 140 LSAQAQVSMVPIFEKGMELVDKMDAEKQNAVGESAGEKATRLTETGVVDVQWGR 199
 Db 137 FTSGKLKEMFPITIAQYGDVLVRLRKE-AEKG-----NPNVNMKDMFGA 178
 QY 200 ATLDMWALAGDYKSDSLQNKTNELVAVFGLTGPAPTLDSP--KAIMMDV-PEFTM 256
 Db 179 YSMDVITGTAFGVNDISL-NPDPFVEHSKNLRLFRP-FDFPILSILLFPNPFVAIL 236
 QY 257 KRRHEIPLTQGLAVSRVRGIELMEQKQAVLGSASDAQVKKVQGVQEDILSLAVRANIAA 316
 Db 237 -----NITLFPKSTVDFTKSVKKI-----KESRLTDKQNRVLDLQLMINSONSK 282
 QY 317 NLPSQKLSDEVLQAISNLLFAGVETSTVLTWFMHRLSDKAVQDKLEECICQIETDM 376
 Db 283 EIDNHKALSDIELVQASTIFIGGYETTSSTLSFIYELTTHPHVQQLQEE---IDATF 339
 QY 377 -----PTLDELNALPYLEAFVKESLRDPPSPYANRECLKDEDFIPLAEPVIGRDSVIN 431

DB 340 PNKAPPTVDALQVMEYLDVMVNETRMFPFIAGRLERVCKDVE-----IH 384
QY 432 EVRTTKGTWMLPLFNINRSKFYGEDAEFRERWLEDVTDLSNS-IEAPYGHQASPTS 490
DB 385 GVTPKGTWLVPLFVJHNNPELWPE-PEBFRERFSKNNKDSINPVYVLPFG-----T 437
QY 491 GPRACFGWRFAVABMAFLFVTLRRVQFEI-----ISHPEYEHITLIIRSP 537
DB 438 GPRNCLGMRFAIMNIKALVRILQNFSEFKPKETQIPLKLYTQGLTQPCQPVILKVV--P 495
QY 538 RIVGREXE 545
DB 496 RGLQFQVE 503

RESULT 15
ID CP3L CALJA STANDARD; PRT: 503 AA.
AC O18993;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 3A21 (EC 1.14.14.1) (CYP11A21) (P450 CW3A-10).
GN CYP3A21.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.
OX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97223367; PubMed=9056237;
RA Igarashi T., Sakuma T., Isogai M., Nagata R., Kanataki T.;
RT "Marmoset liver cytochrome P450s: study for expression and molecular
cloning of their cDNAs".
RL Arch. Biochem. Biophys. 339:85-91(1997).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
tissues by various foreign compounds, including drugs, pesticides,
and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D31921; BAA22156.1; -.
DR HSSP: P14779; LJPZ.
DR InterPro: IPR001128; Cytochrome P450.
DR InterPro: IPR008072; EP450_CYP3A.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR01689; EP450IICYP3A.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 442 442 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 503 AA; 57564 MW; 9BFD421D72C76D6A CRC64;

Query Match 13.2%; Score 377.5; DB 1; Length 503;
Best Local Similarity 27.1%; Pred. No. 1.1e-14;

Matches 148; Conservative 90; Mismatches 184; Indels 125; Gaps 26;
QY 13 LAAPSW--ASIAFSLYLAPRRS-SLYNLQG-PNHTNY-FTGNFLDIISARTG-----E 61
DB 7 LAVETWLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGT---VLYYRQGFWKFDME 63
QY 62 EHAKEYEKYSTURFAGIAGAPVLNSTDEKVNHW-KEAYD-----YKPGMAARVL 113
DB 64 CYKYGKMGMI---YDG--RQPVLAITDPIITKVLVKECVSVFTNRRPFGVGMKSAI 118
QY 114 RIATDGVVTAEGEAHKRRRIMIPSLSAQAVKSVVPIPLEKGMELVDKMMEDAAEKDMA 173
DB 119 SIAQDD-----EMKRIISLLSPFTTSGKLKEMVPIIAQYGEVLVRLRE-AEKG-- 167
QY 174 VGSAGEKKATRLTEGVDVKWVGATLDVMAAGFDYKSDSLQNKTNELVAVFGLTD 233
DB 168 -----KPINMKDIFGAYSDVITGTSFGVNIDSLNNPD----- 201
QY 234 GFAPTLDSFKAIM-WDFV-PYPRTMKRREI-PLTQGLAVSR-----RVGIELMEQ 281
DB 202 ---PFVESTKLLRFDFLOPFFLSITIFPFLFILLEALNISMPFRDSTSLAKSIKRIKE 258
QY 282 KKQAVLGSASDAQVDKXQVQGR--DILSLVLRANIAANLPESQKLSDEVLQAQISNLLPA 339
DB 259 SR-----LKDTHKRVDFQLMIDSONSKETESDKALSDLELVAQSIFIPA 305
QY 340 GYETSSVLTWTFHRLSEDKAVODKLRBEICQIDTDM-----TLDELNALPYLBAFYKE 394
DB 306 GYETTSSTLSFIMYELATHEDVQCKLQEE---IDAVLPNKAATYDVTVLQMEYLDWVNE 362
QY 395 SLRLDPPSPYANRECLKDEDFIPLABPV-GRDGSVINEVRIYKGTVMPLPLNINRSKFI 454
DB 363 TLRLFFPLAMRLERVCKDVE-----INGVFIPKGVVVMIPSYALHYDP-K 406
QY 455 YGEDAEFRERWLEDVTDLSNS-IEAPYGHQASFTSGPRACGWRFAVAEMKAFIVTL 513
DB 407 YWTEPEKFLPERFSKNNKNIDPIYITPFG-----TGNRCIGMRFALMNNKLALIRVL 460
QY 514 RVQFEP 520
DB 461 QNFSKP 467

Search completed: April 2, 2004, 13:59:11
Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 2, 2004, 13:56:29 ; Search time 38.5 Seconds
(without alignments)
4564.770 Million cell updates/sec

Title: US-10-066-007-3

Perfect score: 2852
Sequence: 1 MFILVLTGALGLAARSNAS.....RIVGREKEGYQMLQVKPVE 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	425	14.9	576	3 Q8J2Q5	Q8J2Q5 gibberella
2	419	14.7	593	10 Q9C6S0	Q9C6S0 arabidopsis
3	419	14.7	595	10 Q93VK5	Q93VK5 arabidopsis
4	417	14.6	503	6 Q97689	Q97689 sus scrofa
5	417	14.6	517	10 Q8LIR5	Q8LIR5 oryza sativ
6	403.5	14.1	503	6 Q95L35	Q95L35 sus scrofa
7	403	14.1	515	10 Q9LUD2	Q9LUD2 arabidopsis
8	401	14.1	515	10 Q8W1E1	Q8W1E1 arabidopsis
9	400.5	14.0	519	10 Q9ASR3	Q9ASR3 arabidopsis
10	397.5	13.9	519	10 Q8LHV0	Q8LHV0 oryza sativ
11	394	13.8	503	4 Q8SK3	Q8SK3 homo sapien
12	393.5	13.8	544	10 Q9FE1	Q9FE1 oryza sativ
13	392	13.7	497	11 Q8CJF2	Q8CJF2 rattus norv
14	389	13.6	492	5 Q8ISJ7	Q8ISJ7 mamestra br
15	387	13.6	523	10 Q8LQ77	Q8LQ77 oryza sativ
16	386.5	13.6	511	6 Q8HY6	Q8HY6 capra hircu

17	385.5	13.5	535	4 Q9H241	Q9H241 homo sapien
18	384.5	13.5	531	10 Q9PED3	Q9PED3 oryza sativ
19	382.5	13.4	527	10 Q8LGM8	Q8LGM8 zea mays su
20	382	13.4	430	4 Q7Z448	Q7Z448 homo sapien
21	381	13.4	503	6 Q8HZK1	Q8HZK1 canis famil
22	381	13.4	520	10 Q8LAQ4	Q8LAQ4 oryza sativ
23	380.5	13.3	505	10 Q9LUD0	Q9LUD0 arabidopsis
24	374	13.1	533	5 Q964T2	Q964T2 blattella g
25	371	13.0	507	10 Q8L308	Q8L308 oryza sativ
26	368.5	12.9	523	10 Q9FDZ1	Q9FDZ1 oryza sativ
27	365	12.8	512	10 Q9LUC8	Q9LUC8 arabidopsis
28	365	12.8	520	10 Q48786	Q48786 arabidopsis
29	364.5	12.8	503	5 Q8WQM9	Q8WQM9 anopheles g
30	363.5	12.7	511	10 Q64631	Q64631 arabidopsis
31	363.5	12.7	512	10 Q9LUD3	Q9LUD3 arabidopsis
32	363	12.7	502	11 Q06884	Q06884 rattus norv
33	362.5	12.7	523	10 Q9CAD6	Q9CAD6 arabidopsis
34	362	12.7	506	10 Q9LUD1	Q9LUD1 arabidopsis
35	361	12.7	528	10 Q9FE60	Q9FE60 oryza sativ
36	360	12.6	512	10 Q8W2N3	Q8W2N3 vicia sativ
37	359.5	12.6	512	10 Q9LUC9	Q9LUC9 arabidopsis
38	357.5	12.5	512	10 Q9LUC5	Q9LUC5 arabidopsis
39	357.5	12.5	528	10 Q8L174	Q8L174 zea mays su
40	356.5	12.5	523	10 Q8LAQ9	Q8LAQ9 arabidopsis
41	356	12.5	504	11 Q9EQW4	Q9EQW4 mus musculu
42	355.5	12.5	512	10 Q9ZW95	Q9ZW95 arabidopsis
43	354.5	12.4	506	10 Q7XYT4	Q7XYT4 oryza sativ
44	354.5	12.4	560	10 Q8W372	Q8W372 oryza sativ
45	354.5	12.4	560	10 Q7XD91	Q7XD91 oryza sativ

ALIGNMENTS

RESULT 1
Q8J2Q5 PRELIMINARY; PRT; 576 AA.
ID Q8J2Q5
AC Q8J2Q5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fun15p.
GN Fun15.
OS Gibberella moniliformis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=117187;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M-3125;
RX MEDLINE=93343882; PubMed=10413619;
RA Proctor R.H., Desjardins A.E., Plattner R.D., Hohn T.M.;
RT "A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in Gibberella fujikuroi mating population A.";
RL Fungal Genet. Biol. 27:100-112(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M-3125;
RX MEDLINE=21585628; PubMed=11728154;
RA Seo J.A., Proctor R.H., Plattner R.D.;
RT "Characterization of four clustered and coregulated genes associated with fumonisin biosynthesis in Fusarium verticillioides.";
RL Fungal Genet. Biol. 34:155-165(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=M-3125;
RX Proctor R.H., Brown D.W., Plattner R.D., Desjardins A.E.;
RT "Co-expression of fifteen contiguous genes delineates a fumonisin biosynthetic gene cluster in Gibberella moniliformis.";
RL Fungal Genet. Biol. 0:0-0(2003).
DR EMBL; AF155773; X074818.1;
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.

Query Match 14.9%; Score 425; DB 3; Length 576;
Best Local Similarity 24.9%; Pred. No. 1.4e-22;
Matches 147; Conservative 100; Mismatches 230; Indels 114; Gaps 21;

DR InterPro: IPR001128; Cytochrome_P450.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00067; P450_1.
DR PRINTS: PR00385; P450.
DR DR PROSITE: PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 576 AA; 64918 MW; B1693773CA36A8FB CRC64;

QY 13 LAAFSWASTAFSLYLAPRRSSLYNLQGN--HTNYFTCNFLDILSARTGEHAKYREKY 70
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
50 ISHFWSAI-IWPTFSPRLR-----QLPNVSDGWLSKETLRUVSEGRGVPSQDWINSL 102
QY 71 GS-----TLRPAGIAGAPVLNSTDPKVNHYM-KEADYPKPMAARLVRIATGTGGVTAE 125
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
103 SNRPVDLARYSFGFERLLIIISPKALAEVLTKSYDFRKPLGLIVSELKQTATGMGVILLAE 162
QY 126 GEAAKHHRIMIPSLGAOAVKMVPFIPEKGMLVKMEDAAEKMDAMVGESAGEKKATR 185
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
163 GSEKSQRALKATAFYNRHKNIPLYFWNDVAGEPATVL-----EQKIPTG-----TPTS 212
QY 186 LETEGVDKWDWGRTALDNVALAGFDYKSDSIQNKNELYIAVFGLTDGFAPTLDLS---- 241
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
213 DTTDAVIDVMASRAYLTDIIGRAGMGQGDFATQDDSRHQAYRMI---FFPISRGAIFLA 269
QY 242 -----FKAMWDVPVFRMTKRHBHPITQQGLAVSRVGIEMEOCKQAOLGSASDAQVD 296
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
270 LRLIPIPERLVNWIPIURUKRUH-----GLQVRISKQEILIRERKSKI--KKQKAGVD 321
QY 297 KKDVCQGRDISLLRVANIANLPESOKLDEEVLAQISNLFPAGYTSTVLTMWFHRLS 356
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
322 N---SGNDILTALLNGV-----FTDEQLIDQLTFLAAGHETTATALTWAIIYLC 369
QY 357 EDXAVDKUREBICQ-----IDTDMTFLDELN-----ALPYLEAFKESLRLLDPSP 403
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
370 KOPEVQNRLREEIRMFHPNPKGWPRSERSSNTLOQAIDFKLPYLNVCLEVWRVFAPIP 429
QY 404 YANRECLKBEDFIPLAEPVIGRDGVSINEVRITKTGMTVMPLPFNNINRSKFYIGEDAEFR 463
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
430 LTWTREATCD-----TWILHTFPAGTRIILARPVTNRDSALWGPDANNFN 474
QY 464 PERKLEDVDTSLNSIBAPVGHQ-----ASFISGPACFGWRFAVAEMKAF 508
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
475 PDR-----NKORDSTAAPVTOBEVRGRTEARSNVADTLFHGPSRCISGCQSFAERVEFAIL 529
QY 509 LFVTLRAVQPPIISHPEVEHIITLSR---PRIVGREKEGYQMRLQVKPV 556
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
530 LATLIANFEQ-QIEDESLLDERNISIRSATSRIVG-----GLKVVRPI 573

RESULT 2
Q9C6S0 PRELIMINARY; PRT; 593 AA.

ID Q9C6S0
AC Q9C6S0
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome p450, putative.
EN F5M6.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=Cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen R.F., Chin C.W.,
RA

Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Eggu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.P., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A., Millischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pai G., Peterson J.G., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vyatskaja V.S., Walker M., Wu D., Yu G., Fraser C.N., Venter J.C., Davis R.W.; "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana." Nature 408:816-820(2000).

-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
ENBL; AC079041; AAG50718.1; --
DR FIR; F86441; F86441.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_p450.
DR Pfam: PF00067; P450_1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 593 AA; 66643 MW; E80CE9B8B2BD199 CRC64;

Query Match 14.7%; Score 419; DB 10; Length 593;
Best Local Similarity 26.4%; Pred. No. 4.1e-22;
Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NFILDLSARTG--EFHAKYREKYGST-----LFAGI-----AGAPVLN 86
DB NVLDEFMFDTGDDQYPKYPEAKGISQAVRNAEFFIPLYELFYLGTRITFGFKSFLI 152
QY 87 STDPKVNHVME-KADVYPKPMARVLRIATGTGWTAEGEAHKKRHRIIMPISLAQAV 145
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
153 VSDPSIAKHILKONAKAYSK-GILAIEDFDVWGKLIPADGEIWRERRRAIVPALHMQYV 211
QY 146 KSNVPIFLEGKLMELVDKMEMDAEEKDMAVGESAGEKKATLETGEDVDKVVGRATLDWM 205
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
212 AMISLFGASDLCCQL--DAA-----ALKGEVEMESLSRLTLDII 253
QY 206 ALAGPYKSDSLQNTNELNYIAFVGLTDGPAITLDSFKAIWMDVPVPTMKRHEIPT 265
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
254 GKAVFNYPDSTLNTDGIEAVTYTLREAERSVIP--VMD-IPIWKDIS-----PRQ 305
QY 266 QGLAVSRRVGIELMEQKQAVLGASDAOAVDKDVQ-----GRD--ILSLVRANIAA 316
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
306 RKVATSLKINDTLD-----LIACTKRWVEEBELQFHEYMNREDPSILHLF----- 353
QY 317 NPESOKLSDEEVLAQISNLLPAGYTSTSVLTWNFHRLSDKAQVDKLBREIQCIDTD- 375
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
354 -LASGDVSSKQLRDDLTMLTAGHETSAAVLTWTFYLLTPESPVAKLOBEVDSVGDR 412
QY 376 MPLTDELNALPYLEARFKESLRDPSPVANRECULKDEDIFIPLAEPVIGRGDSVINERI 435
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
413 FFTIQMKKLKLTTRVMNESRLYIQPPVLIRSI-----DNDILGEYPI 457
QY 436 TKGTVMYMLPLFINRSKFYIGEDAEEFERPWLEVDTD-----SUNSIEAPYHGHAQSFISG 491
DB KRGEIDIFISVNMHLRSP-LHWDDAEKFNPERWPFGDGNPETNQNFSLYLPFG-----GG 510
QY 492 PRACFGWRFAVAEMKAFVTLERVQPE 519
DB : : : : : : : : : : : : : : : : : : : : : : : : : :
511 PKFCIGMPASPENVVAVIAIWLRFRNFQ 538

RESULT 3
Q93VK5 PRELIMINARY; PRT; 595 AA.
ID Q93VK5
AC Q93VK5;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE At1931800/68069.m00159.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota, Viridiplantae, Streptophyta, Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 ET "Arabidopsis cDNA clones";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 ET "Arabidopsis cDNA clones";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
 RA Yamada K., Yu G., Yuan S., Davis R.W., Theologis A., Ecker J.R.;
 ET "Arabidopsis ORF clones";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AY058173; AAL25587.1; -;
 DR EMBL; AY056446; AAL08302.1; -;
 DR EMBL; AY12017; AAM98281.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 595 AA; 66845 MW; A548A33A48323B7D CRC64;

Query Match 14.7%; Score 419; DB 10; Length 595;
 Best Local Similarity 26.4%; Pred. No. 4.1e-22;
 Matches 134; Conservative 95; Mismatches 179; Indels 100; Gaps 19;

QY 50 NPLDILSARTG--BEHAKYREKYST-----LRFAGI-----AGAPVLN 86
 DB 95 NVLDFMFWTQDQDPKVPKAGSIQVRNEAFIPLEYLFTYGIPLTGTGKSFLL 154
 QY 87 STDPKVFNVHMK-AYDYKFGMAARVLRIATGQVVTAEAGAKHRRHRIPIFSLAQAV 145
 DB 155 VSDPSIAKHILKONAKYSK-GILAEILDVFMGKLIPADGEIWRRRRAIVPALHOKV 213
 QY 146 KSNVPIFEKGMELVDKMDAAEKDVAAGSAGEKATRLTEGVVDKDWGRATLDVM 205
 DB 214 AAMISLFGASDRLCQKL--DAA-----ALKGEVEMESLFGRLTLDII 255

QY 206 ALAGPDYKSDSLQNKTNELNYAVFGLTDPGAPTLDSFKAIMWDFVPVYFRTMKRRHIEPLT 265
 DB 256 GKAVFNFDSTNDTNGTGWIEAVITVLRADRSVSPFP--VWD-IPWKDIS-----PRQ 307
 QY 266 QGLAVSRVRGIELMEQKQKQAVLGSASDAQVDKDVQ-----GRD--LISLLVRANIAA 316
 DB 308 RKVATSLKINDLDD-----LIATCKMVEEELQFHEEYMERDPSILHFL----- 355
 QY 317 NLPESQKLSDEVLQAQINLLFAGYETSTVLTWMFHLSEDKAVQDKLREICQIDTD- 375
 DB 356 -LASGDVSSKQDRDLMTMLTLAGHETSAALTTFTFYLLTTPSVVAKLQEDSVIGDR 414
 QY 376 MPTLDLINALPYLEAFVKESRLDPPSYANRECLKDEDFLAPVIGRDSVINEVRI 435
 DB 415 FPTIQDKKLKYTRVMNESRLYQPVPVLRRSI-----DNDILGEVPI 459
 QY 436 TKGTMVLPFLNINRSKFIYGEDAEFPERWLEVD-----SLNSIEAPYGHQASFI 491
 DB 460 KRGEIDIFISVWNLHRSP-LHWDDAEKFNPERWFLDGPMPNETNQNFSLYPFG-----GG 512
 QY 492 PRACFGMREFAEMKAFVLTLLRRVQFE 519
 DB 513 PRKICGMFASFNVAIAMLIRRFNFQ 540

RESULT 4
 O97689 PRELIMINARY; PRT; 503 AA.
 AC O97689;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cytochrome P450.
 OS Sus scrofa (Pig).
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 [1]
 SEQUENCE FROM N.A.
 RA Schuetz E., Andag R., Wieland E., Oellerich M.;
 RA "Porcine Cytochrome-P450 with high protein homology to human CYP3A4
 (Nifedipine oxidase J04449).";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF109068; AAD04628.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR008072; EP450_CYP3A.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01689; EP450IICYP3A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 503 AA; 57267 MW; 22DB19F43C5988BB CRC64;

Query Match 14.6%; Score 417; DB 6; Length 503;
 Best Local Similarity 27.9%; Pred. No. 4.4e-22;
 Matches 160; Conservative 82; Mismatches 211; Indels 120; Gaps 24;

QY 12 GLAAFSWASIA--FFSLYLAPRS-----SLYNLQGNHTNYFTGNFLDILSARTGEEH-- 63
 DB 6 GFSFTETVLLATSLVLLYLYGTYSHGLFKKLGIPGRPLPYF-GN---ILGVRKGVDFPD 61
 QY 64 AKYREKVGSTLRFAGTAGAPVLNSTDPKVFNVHM-KEADY-----PKFGMAARVL 114
 DB 62 KKCFQYQKGMWGFDD-GRQPVLAITDPMIKTVLKECYSVFTNRRSFGFRGAMRT---- 116
 QY 115 IATGDGVVTVEGAHKHRRIMIPSLSAQAVSKSMVPIFEKGMELVDKMDAAEKDVA 174
 DB 117 -----ALSLADESEWKIRTLISFTTSGLKEMFFLIISHYGDLLVSNLRKS-AEK 167

175 GBSAGEKATRLTETEGVDYKDWGVRATLDVNLALAGFDYKSDSL-----QNKTNELVY 226
 168 -----KPVTKMDIFGAYSDVITSTAFGVNTDFLNNPDPPFVNSKLLKF 213
 227 AFVGLTDFGAPLDSFKAIMWDFVYFRTMKRHEIPLTOGLAYS-----RVGIELMBQKK 283
 214 SP-----FSLII-----FPFL-----TPILEVNTLTPKSVNFFWRSIK 254
 284 QAVLGSADQAVKDVQGRDILSVLRANIAANLPSQKLSDEVLQAQISNLLFAGYET 343
 255 R-----MKESRLKQTHRVDFLQLMINSQNSKETDTHKGLSDEELVAQGVFFIAGYET 309
 344 SSTVLTWPHRLSDEKXAVODKLEELICQI---DTQPTLDELNALPYLEAFVKESLRIDPP 401
 310 TSSSLLYELATHPDVQCKLQSEIDATFFSKALPSFDALAEYEDLVNNEILRLYPI 369
 402 SPYANRECLDKEDFIPLAEFVIGRDSVINEVRITKGTVMVPLFNINRSKFIYGEDABE 461
 370 AALERVCKDVE-----LHGVSVKGTVMVVPFVSHRDBELWPE-PEE 413
 462 FRPERWLEDVTSLSNIEAPYGHQASFISSGRACGWFVAEMKAEFLVTLRVQFEPI 521
 414 FRPERFSKQKDSIN---PTY-LPFGTGPNCIGMRPALNMKLLALVRVLQNSFKPC 468
 522 -----ISHPEYEHITLIISRPVIG 541
 469 KETQPLKLSQGLIOPEKPIILKVPRDGTGV 501
 RESULT 5
 Q8LIR5 PRELIMINARY; PRT; 517 AA.
 AC Q8LIR5;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 GN OJ1332.C12.12.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OK NCBI_TaxID=39947;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 clone:OJ1332.C12.12";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AP003752; BAC10039.1; -;
 DR Gramene; O8LIR5; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 517 AA; 58038 MW; 3C3B35745026E138 CRC64;

Query Match 14.6%; Score 417; DB 10; Length 517;
 Best Local Similarity 26.5%; Pred. No. 4.6e-22;
 Matches 154; Conservative 99; Mismatches 202; Indels 130; Gaps 24;
 6 LITGALGLAFAFSAFFSLYLAP-----RRSLYNLQGNHNTYFTGNFLDILSARTGE 61
 7 MYAAVAANVASAFDAVVKLVWRPRAITRLRAQGVGGPGY-RPFGNGLGEI--RLRD 63
 62 EHA-----KYREKYGSTURFAGIAGAPVLNSTDKPVNHYMKAY 101
 64 EGAGVLDVSHDFVPIVQPHFRKWIPLYKGTMY-WFGARPTICLADVSVWRVQLSDRT 122

102 D-YPK-----PGWAARVLRIATGCVTAEGEAHKKHREIMIPSLSAQAVKSMVPIFLEKCM 157
 123 GMYKPNVSNPYFARLL-----GKGLVLTGDGWRKRRKRVHPAFNMDKLCMT----- 170
 158 ELVDKMWEDAAEKDMVAGESAGEKKATRLTEGVDYKDWVGR-----ATLDVNLALAGF--D 211
 171 -----VTMSDCAQSMISEWES-----ELGTGDIVELSRFPELTADVISHTAFGSS 219
 212 YKSD-----SLQNTKNELYAFVGLTDFGAPTLDSFKAIMWDFVYFRTMKRHEIP 263
 220 YKEGQVFLAQREIQ-----FLAF-----STFLSI-----QIP 247
 264 LTQGLAVSRVRVGIEMBOKKQAVLGSASDAQVKKQV---QGRDILSVLRANIAANLPS 321
 248 GSSVLPYTKCNLTKWSVDKVRSMITDIKSLRNKNDVAGYGNLGLMLEA-CAPSHGES 306
 322 Q-KLSDEVLQAQISNLLFAGYETSSVLTWPHRLSDEKXAVODKLEELIC-QITDMPIL 379
 307 QPOLSMDBIIAECKTFFAGHDTTSHLLTWTMFLSTHPWQEKLEEVATECDGKVPFG 366
 380 DELNALPYLEAFVKESLRLLDPPSPYANRECLDKEDFIPLAEFVIGRDSVINEVRITKGT 439
 367 DMLNKLKLVNMFLETLRLYGPVAFIQRVNAELE-----LGGITVPEGT 411
 440 VMVPLFNINRSKFIYGEDAEFRPERWLEDVTSLSNIEAPYGHQASFISSGRACGWF 499
 412 VLSIPTIATHRDKEVGEDADIFKPERFKNGSKAGKYPNA-----LISFSSGPRACIGN 467
 500 FAVAEKMAFLVTLRVQFEPIISHPEYEHITLIISRP 538
 468 FANTEAKAVIAMILQRFSP---TLPKIVHVPTDITLRPK 505
 RESULT 6
 Q9SL35 PRELIMINARY; PRT; 503 AA.
 AC Q9SL35;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Cytochrome P450 3A.
 GN CYP3A.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OK NCBI_TaxID=9823;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Soucek P., Zuber R., Anzenbacherova E., Linka M., Anzenbacher P.;
 RT "CDNA of minipig cytochrome P450 3A";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF24780; AAL13316.1; -;
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. . .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR008072; EP450_CYP3A.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01689; EP450IICYP3A.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 503 AA; 57085 MW; 5A17128CA9E50914 CRC64;

Query Match 14.1%; Score 403.5; DB 6; Length 503;
 Best Local Similarity 27.8%; Pred. No. 4.4e-21;
 Matches 152; Conservative 81; Mismatches 191; Indels 123; Gaps 22;
 12 GLAAPSASIA--PFSLYLAPRRS-----SLYNLQGNHNTYFTGNFLDILSARTGEEH-- 63
 6 GPSTEWLVLLSLVLLLYGTYSHGLFKLIGIPQRPPLPYF-GN---ILGRKGVDFHD 61


```
QY 218 ONKTNELVAVFVGLTGDPAFTLDFKALMMDVFPYFTMKRRHEIPLTQGLAVSRVRGIE 277
Db 222 -----SSYVEGIEVFRSQM-ELKRCYTTSLNQVSIPTQYLPPTPSNIRVM 265
QY 278 LMEQKQAVLGASDAQVKKDVOGRDITLSLVANIANANLPSOKLSDEVLAIQISNLL 337
Db 266 KLERKWDNSIKRITSSRLQSKSDYGDLLGLLKAYNTEG--KERRKVSIEIIEHCRTFF 323
QY 338 FAGYETSITVTWTFHRLSEDAKAVQDKLREEI---CQIDTMDPTLDELNALPYLEAFVKE 394
Db 324 FGHETTENLLAWITMLLSLHQDQWQEKLEBIFKECGKE-KTDSSETFSKLKMNVMVME 382
QY 395 SURLDPPSPYANRECKLDEDFIPLAEPVIGRGSV---INEVITKGTWMLPLFNINRS 451
Db 383 SURLTGP-----VSALAREASVNIKUGDEIPKTTVVIPLLKQHS 424
QY 452 KFIYGEDAEFRPERLWSDVTDLSNIEAPYGHQASFTSGPRACFGWRFVAEMKAPLFV 511
Db 425 KTLWGSADAKFMPFANGVSRANHPNA---LLAFSGVPRACIGQNFVMEIAKTVLTM 480
QY 512 TLRRVQFPPIHSHPIEH--TLIISRPRIVGREKEGYQMLQVKE 557
Db 481 ILQRFPRFISLQD--EYKHTPDVNTIQD-----YGFVMLQPLE 518

RESULT 10
Q8LHV0 PRELIMINARY; PRT; 519 AA.
AC Q8LHV0;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN P0025D09.13.
OS Cryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0025D09."
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR ENBL; AF004264; BAC10362.1; -.
DR Gramene; Q8LHV0; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 519 AA; 59547 MW; 77E79E9C1428FE38 CRC64;

Query Match 13.9%; Score 397.5; DB 10; Length 519;
Best Local Similarity 26.0%; Pred. No. 1.3e-20;
Matches 155; Conservative 80; Mismatches 225; Indels 137; Gaps 22;

QY 10 ALGLAIF-----SWASIAFSLYLAP-----RSSLNLOGPNHTNYFTGNFLDI 54
Db 2 AMGLLWMAVAAAAAASVHLVWREPRALSRLRAQGVGGY-RFFSGNLGEI 60
QY 55 LSAR-----TGEHAKYREKYGSLRFAGIAPVLNSTDPKVFNVH 96
Db 61 KRFRGAGVNLVNSHDFLPVQPHFRKWIPLYGTFLY-WFGQPNICLADYVMWQV 119
QY 97 MKEAYD-YPKPGMAARVLRIATGDGVVTAEGBAHGRHRIMIPSLSAQAVKSWVDFLEK 155
```

```
Db 120 LSRDTGIYPKNTNPHFVRL-LGKGLVLTDGDEWKEHRRKVVHPAFNMCKLKWMT----- 172
QY 156 GMELVDQMEDAAEKQAVGESAGEKKAFLTEGVDVKDWVGRATLDMALAGF--DYK 213
Db 173 -----MTNSDCGRSMWSESELAAGGLVE---IELSRFEELTADVISHTAFGSSYK 223
QY 214 SDS---LQKNTNELYVAFVGLT---DGFA--PTLDSFKAIMMDVFPYFTM-----KRHH 260
Db 224 EGKQVFLAQRELQFLAFSTFLTQVIGFSVLPIMKFKT--WSLDKKVRGMLMDIILKTRH 281
QY 261 EIPLTQGLAVSRVRVGIEMEQKQAVLGASDAQVKKDQV--QGRDILSLVANIANL 318
Db 282 A-----NKDVAGYGNLGLGMLLEACAPHG 306
QY 319 PESOKLDEEVLAIQISNLLFAGYETSTVLTWTFHRLSEDAKAVQDKLREEIC-QIDTMDP 377
Db 307 ESCQLSDEIIECKTFFPAGHDTTSHLLTWTFLLSTHDPWQEKLEBEIAECGDKVP 366
QY 378 TLDELNALPYLEAFVKESSLRDLPPSPYANRECKLQSDDFIPLAEPVIGRGSVINEVRIYK 437
Db 367 TGDMLNKLKVMNPLLETLRLYSPVSLIRKVDTDIE-----LGGIKMPE 411
QY 438 GTWMLPLFNINRSKFIYGEDAEFRPERLWSDVTDLSNIEAPYGHQASFTSGPRACFG 497
Db 412 GALLTIPIATHIRDKVEWGEDADEFRPERFENGVTAAKHNA-----LLSFSSGPRSCIG 467
QY 498 WRFVAEMKAPLFVTLRRVQF--EPIISHPEYEHITLIISRPRIVGREKEGYQMLQ 552
Db 468 QNFAMIEAKAVIAMILQRFSTLSPKYVHAPTQVITL-----RPKYGLPMILK 515

RESULT 11
Q86SK3 PRELIMINARY; PRT; 503 AA.
AC Q86SK3;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Cytochrome P450.
GN CYP3A43/CYP3A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=21839017; PubMed=11726664;
RA Finta C., Zaphiropoulos P.G.;
RT "Intergenic mRNA molecules resulting from trans-splicing."
RL J. Biol. Chem. 277:5882-5890(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Zaphiropoulos P.G.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR GO; GO:0006712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008072; EP450_CYP3A.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01689; EP450IICYP3A.
DR PROSITE; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 503 AA; 57394 MW; 4141F95B28B34A5E CRC64;

Query Match 13.8%; Score 394; DB 4; Length 503;
Best Local Similarity 28.7%; Pred. No. 2.2e-20;
Matches 142; Conservative 83; Mismatches 177; Indels 92; Gaps 20;

QY 47 FTGNFLDILSARTG-----EEHAKYREKYGSLRFAGIAPVLNSTDPKVFNVH-K 99
```

Db 46 FLGN---ILSYHKGFCMEFMECHKYGVKMGF---YDG--QQPVIAITDDPMIKTVLKE 97
 QY 100 AYDY---PKPGMAARVLIATGDGVVTAAGBAHKEHRRIMIPSLSAQAVKSWVPFLK 156
 Db 98 CYSVFTNRFPFGVGMKSA---ISIADEBWKRLSLLSPTTSGLKEMVPIIAQYG 153
 QY 157 MELVDKMDAEDAEKMAVGESAGEKATRLT-EGVDVQKDWVGRATLDVWALAGFDYKSD 215
 Db 154 DVLVRNLREA-----ETGKRVTLKDVFGAYSDVITSTSGFNID 194
 QY 216 SLQNKTNELVYAFVGLTDCGFAPTLDSFKAIM-WDFV-PYFRIMK-RRHEIPTLTOGLAV-- 270
 Db 195 SLNPNPD-----PFVENTKLLRFDFLDPFPLSITVFPFLIPILVNICV 240
 QY 271 -SRVIGIEMEKQKQAVLGASDAQVDKDVQGRDILSLVRAIANANLPSQKLSDEEV 329
 Db 241 FPREVTNFRKSVKR-----MKESLDEPTQKRVDFLQIMDSQNSKETESHKALSLEL 295
 QY 330 LAQISNLLFAGYETSTVLTWTFHRLSEDAVQDKLREEICQI--DTDMPTLDELNALPY 387
 Db 296 VAQSIIFIFAGYETSSVLSFMYELATHPDVQOKLQEIIDAVLENKAPPTVDVLQWEY 355
 QY 388 LEAFVKSRLDPPSPYANRECLKDEDPIPLAEPVIGRDSVINEVRITKGTWMLPLFN 447
 Db 356 LDMVNETLRLPFIAMRLERVCKDVE-----INGMFPKGVVWVIPSYA 400
 QY 448 INRSKIYGEDAEERPERWLEDVTDLSNS-TEAPYGHQASFISCPACFCWRFAVEMK 506
 Db 401 LHDP-KYVTEPEKFLPERFSKKNKNDIPYITTEFG-----SGPRNCIGMRPALNMK 453
 QY 507 AFLFVTLRRVQEP 520
 Db 454 LALIRVLQNPSEKP 467

RESULT 12
 Q9FEEL ID Q9FEEL PRELIMINARY; PRT; 544 AA.
 AC Q9FEEL;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative cytochrome P450.
 GN P0688A04.9 OR P0006C01.24.
 OS Oryza sativa (rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone:P0688A04.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone:P0006C01.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF002839; BAB19103.1;
 DR EMBL; AF002744; BAB19082.1;
 DR Gramene; Q9FEEL;
 DR GO; GO:000497; F:monooxygenase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450;
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.

KW Heme; Monooxygenase; Oxidoreductase
 SQ SEQUENCE 544 AA; 60867 MW; 273EAF5968D1A024 CRC64;
 Query Match 13.8%; Score 393.5; DB 10; Length 544;
 Best Local Similarity 26.4%; Pred. No. 2.7e-20;
 Matches 159; Conservative 83; Mismatches 215; Indels 145; Gaps 25;
 QY 6 LLTGALGLAASWASIAFFSLSYLAPRR-SSLYNLQGNHNY-----FTGNFLDLSAR 58
 Db 31 LLGGVAAALLLVWAAQMLEWAWLAPRMERALRAQGLRGTOYRFLGDLTDLRLVTAAR 90
 QY 59 TG-----EEHAKYREKYGSLRFPAGIAGAPVNLSTDPKVFNVHV 96
 Db 91 SKPVMDRPHDFIPRVAPLLHRALEHGR-----VSFTWFGMPRVTTTDPDLVREV 142
 QY 97 MKEAY-DYFKGMAARVLIATGDGVVTAAGBAHKEHRRIMIPSLSAQAVKSWVPFLK 155
 Db 143 LSNFGHEFKTKATRLSKLLVG-GLVLHGEKWKVRRHRRIMNFAHAEKLRMLPAPSAS 201
 QY 156 GWELVDKMDAEDAEKMAVGESAGEKATRLTETEGVDVQKDWVGRATL--DWMALAGFDYK 213
 Db 202 CSELIGRW-----ENAVAASVGKAE-----LDI--WPDFQNLSDGVISRAAFGVR 244
 QY 214 SDS-----LQNKTNELVYAFVGLTDCGFAPTLDSFKAIMWDFVYF-----RTMKRHEI 262
 Db 245 RHGEGQIFLLQAEQAEHLV-----QSFRS---NYIPGLSYENNRNK-----283
 QY 263 PLTQGLAVSRV-----GIELMEQKQAVLGASDAQVDKDVQGRDILSLVRAIANAN 317
 Db 284 -----AIDRIKILRGI-IEKQKATKNGAS-----KD---DLLGLLQSNMDYY 326
 QY 318 LPESOKLSD-----BEVLAQISNLLFAGYETSTVLTWTFHRLSEDAVQDKLREEICQI- 372
 Db 327 SDEGKSKGTMVEEIIIDECKLFYFAGMETTAVLLTWTWVALSNHPEWQDRAREEILQVF 386
 QY 373 DTDMPITLDELNALPYLEAFVKSRLDPPSPYANRECLKDEDPIPLAEPVIGRDSVINE 432
 Db 387 GRNKPDIINGVSRKLVTVMLHEVLRLLYFPVVMQRRTYKIE-----LGG 431
 QY 433 VRITKGTWMLPLFNINRSKFIYGEDAEERPERWLEDVTDLSNSLEAPYGHQASFISGP 492
 Db 432 VRYPAGVMSLSPLVFIHRDAAAAGHDAGEFDPGRFAEGVARACKDPCA--GAFPPFSWGP 489
 QY 493 RACFGWREAVAEKAFPLVTLRRVQEB--PIISPEYEHITLIIISRPVIGREKEGYQMR 550
 Db 490 RICIQNFALLEAKVALGNILQRFAPLSPALAPYVTL-----HPQHGVPVR 540
 QY 551 LQ 552
 Db 541 LR 542

RESULT 13
 Q8CJF2 ID Q8CJF2 PRELIMINARY; PRT; 497 AA.
 AC Q8CJF2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cytochrome P450 3A.
 GN CYP3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Matsubara T., Nagata K., Yamazoe Y.;
 RT "Isolation and characterization of a novel rat CYP3A form.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB084894; BAC23085.1;
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.

```

DR GO: 0006118; P: electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 497 AA; 57009 MW; 8C034C25C476F452 CRC64;

Query Match      13.7%; Score 392; DB 11; Length 497;
Best Local Similarity 26.4%; Pred. No. 3e-20;
Matches 140; Conservative 81; Mismatches 176; Indels 134; Gaps 20;

QY 35 LYNLOGPHNTY-----PTGNDILSARTG----BEHAKYREKYGSTLRAG 78
DB 22 LYLGTSTHGNFKLIGSGPKLPFVGN---ILAYRHGFWEFRHC--HKYGDINGFYE 76
QY 79 TAGAPVLNSTPKVFNHVM-KEAY-----DYPKPGMAARVLRATGCGVWTAEGEAHK 130
DB 77 -GRQILAITDPDIKTVLVEKCYSTFNRRSFGPAGILKKAITL-----SEDEWK 127
QY 131 RHRIMPSLSAQAVKSMVPIFEKGMELVDKMDAEDAAEKDMAGVESAGEKATRLTEG 190
DB 128 RLRLTSLTFTSGKLKEMFPI-INQYADLLVKNVKEAEKG-----NP 169
QY 191 VDVKDWGVRATLDVWALAGEDYKSDSLONKTN-----ELYVAFVGLTDGF 235
DB 170 ITMKDIFGAYSMGVITGTSFGVNDLNNPONPFQVKVKKLLKFNFLDPFLSVLPFL 229
QY 236 APTLDSFKAIMW--DFVYFRTMKRRHEIPLTQGLAVSRVRVIGIELMEQKQAVLGASDQ 293
DB 230 TVPFEAFDITVPFKDVKMFRT-----SVERMKENR----- 260
QY 294 AVDKDVQGRDILSILVRANANLPSQKLSDEVLQIISNLFAGYETSTVLTWVFH 353
DB 261 -MOEKVKQBLDFQLMINSQSGGKXESHQGTDEIVEAQSFPIFAGYETSSALSPLY 319
QY 354 RLSEDKAVQDKLREICQIDTMP-----TLDELNALPYLFAFKESLRDPPSPYANRE 408
DB 320 LLATHPDQLKQDE--IDAALPNKAPVTVYDVLVEMBYLDMNLNLTLPFVGGRLERV 376
QY 409 CLKDSDPFLAEPVIGRDSVINEVRIKTMVLMPLFNINSKFIYGEDAESEFRPERWL 468
DB 377 CRKQVE-----INGVIFPKTVVWVPTFALHKDKPCWPE-PEEFCPEFR 420
QY 469 EDVTSLSNS-IEAPYCHOASFTSGPRACFGWRFAVEMKAFVLTLRVQF 518
DB 421 KKNQDSINPYIILPFG-----NGPRNCIGMRPALMNNKIALVRVLQNFSF 465

RESULT 14
Q8ISJ7 PRELIMINARY; PRT; 492 AA.
AC Q8ISJ7
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Cytochrome P450 CYP414.
GN CYP414.
OS Mamestra brassicae.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Hadeninae; Mamestra.
OX NCBI_TaxID=55057;
RN [1]
RP SEQUENCE FROM N.A.

TISSUE=Antenna.
MEDLINE=21956093; PubMed=1200647;
RA Maibeche-Colsen M., Jacquin-Joly E., Francois M.C.,
RA Nagnan-Le Meillour P.;
RT "cDNA cloning of biotransformation enzymes belonging to the cytochrome
RT P450 family in the antennae of the noctuid moth Mamestra brassicae.";
RL Insect Mol. Biol. 11:273-281(2002).

DR EMEL; AY063501; AAL48300.1; -.
DR GO: 0006118; P: electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
SQ SEQUENCE 492 AA; 56219 MW; 4F69897198D9D01B CRC64;

Query Match      13.6%; Score 389; DB 5; Length 492;
Best Local Similarity 22.9%; Pred. No. 5e-20;
Matches 139; Conservative 105; Mismatches 182; Indels 180; Gaps 24;

QY 1 MFILVLTGALGLAAFSWASIAFSLYLAPRSSLYNLOGPN-----HTNYF---TGNF 51
DB 6 VFVVVLA-----LLVSWISLV-----RVSRRENVPGPSPLPLVGNALHFLVVKSEF 52
QY 52 LDILSATGEHAKYREKYGSTLRFAGIA-----GAPVLNSTDPKVFNVHMKEA 100
DB 53 LNLVQ-----RLSEKYGHAFRVHFFSTFYVVICHSKYAEPLVSSTB-----HITK-- 97
QY 101 YDYPKPGMAARVLRATGCGVWTAEGEAHRRHRIMPSLSAQAVKSMVPIFEKGMELV 160
DB 98 -----GRSFSFLTCMLGQGLLTATQGRWKSHERKELTPAFHFNILQNLFLPVFCNQRILT 151
QY 161 DKM--MDAAEKDMAGVESAGEKATRLTEGVDVKWVGRATLDVWALAGEDYKSDSL- 217
DB 152 EKIRGMADGRPIDM-----FPIIALAALDNVNESIM 182
QY 218 -----CNKTNELVAFVGLTDGFAPLDSFKAIMWDFVYFRTMKRRHEIPLTQGLAVS 271
DB 193 GVCMDAQKHSQSEV-----KSIELSIV-----TM--RMQIPFGEDAIF 222
QY 272 RRVGIELMEQKQAVLGASDAQV-----DKDVQGRD-----ILSLVRA 312
DB 223 NLLPYTKQDKALKVLRGQTKVKNARRAELKKNANITLNDSSDICTKNKHTFLDLLLA 282
QY 313 NIANLPSQKLSDEVLQIISNLFAGYETSTVLTWVFHRLSEDKAVQDKLREI--- 369
DB 283 EI-----DGKIDDESVEEDVTFWFEHDTITTSIGVYITLHCLSKRDVQEKIYELKTI 337
QY 370 --CQIDTMPDLDELNALPYLFAFKESLRDPPSPYANRECLKDSDPFLAEPVIGRQD 427
DB 338 FGSEIHRD-PYHLELQMKYLELVKESMLRFPFVPLIERRIMKDCE----- 383
QY 428 SVINEVRITKTMVLMPLFNINSKFIYGEDAESEFRPERWLEDVTSLSNSIEAPYCHOAS 487
DB 384 --VGSGLKLVKGTSTVVMNIFQIQOPDLF-DDPLEFRPERF-----EAPLKNPFS 429
QY 488 ---FTSGPRACFGWRFAVEMKAFVLTLRVQFEPITISHPEYEHITLIISRPRIVGREK 544
DB 430 WLAFSAGPRNCIGQKPAWMLKTIISIVKNFFILPAEEPELS-----ADLVLRSK 481
QY 545 EGYQNR 550
DB 482 NGVNVK 487

RESULT 15
Q8LQ77 PRELIMINARY; PRT; 523 AA.
AC Q8LQ77
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
GN B1051E10.19.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=cv. Nipponbare;
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone: B1051E10."
CC Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AP003764; BAB93411.1; -
DR Gramene; C81Q77; -
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 523 AA; 58192 MW; D5B9C6C0F0965798 CRC64;

Query Match 13.6%; Score 387; DB 10; Length 523;
Best Local Similarity 24.8%; Pred. No. 7.7e-20;
Matches 143; Conservative 94; Mismatches 245; Indels 94; Gaps 19;

QY 3 ILVLTGAL---GLAAPSWSIAFFSLYLAPRSSLYNLOGNHNHYFTGNELDILSART 59
DB 10 ILVVLVSRLLTSALVHLWKPYPATKLFGR-----QGITGPKY-RLFVGSLSPEIKRMKA 62
QY 60 G-----EEHAKYREKYGST-LRFAGIAGAPVLNSTDPKVFHVMK 98
DB 63 AAADEVAAGASHSDFTPIVLPQHSKWATDHGKTFVWLGAVPAVSLGRVE-QVKQVLE 121
QY 99 BAYDYPKPGMAVRVLRATGCVVTAEGEAHKEHRRIMIPSLSAQAQVKSVMVIFLEKME 158
DB 122 RTGSFTKNYMNAN-LEALLGKGLIANGEDWERHKKVHPAFNHDKLFKMSVYMAESVES 180
QY 159 LVDXNMEDEAKONAVGESAGEKATRETEGVVDVVDWVGRATLDVWALAGFDYKSDSLQ 218
DB 181 MVQRWQSQI-----QQAGNQVE-----LDLSRELSELTSVITRSAP----GSSH 222
QY 219 NKTNELYVAFVGLTD-GFAPTLDSFKAIMWDFVYPRTMKRRHEIPLTOGLAVSRVGLIE 277
DB 223 BEGKEVYQAQKELQELAFSSSLDVPALVFLRKLFIIRGNTRAHQ-----VKKSRTMLME 277
QY 278 LMEOKKQAVLGSASDAQVKKDQGRDILSLVYRANIAANLPESOKLSDEEVLAQISNLL 337
DB 278 IIEGRLAKEAABAG-----YGSDDLGLMLEARALEREENGLVLTTOEIIDECKTF 329
QY 338 FAGYETSTVLTWMFHLSEDKAVQDKLBEICQIDTD-NPTLDELNALPYLEAFVKESL 396
DB 330 FAGQDTTSHLVWMTMFLSSNAQWQDKLREEVLTVCQDAIPTDPMANRLKLVNMVLES 389
QY 397 RLDPSPYANRECLKDEDFPLAEPVIGRDSVINEVRITKGTWVMLPLFNINRSKFIYG 456
DB 390 RLYSPVVIIRIAGSDID-----LGNLKI PKGVLSIPIAKIHRDRDWG 434
QY 457 EDAAEFPRERWLEDVTDLSNLSIEAPYGHQASFISGPRACFGMRFAVAEMKAFVTLRRV 516
DB 435 PDADFNFPARFKNGVSRASYPNA-----LLSFSQGPREGCIGQTFAMLESQIAIAMILQRF 490
QY 517 QF--EPIISHPEYEHITLIISRP-----IVGREKEG 546
DB 491 EFRUSPSYVHAPMEAITL---RPRFGLFVLRNLQ 523



STIC SEARCH RESULT FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 04:33:00 ; Search time 170 Seconds
(without alignments)
6306.854 Million cell updates/sec

Title: US-10-066-007-2
Perfect score: 1932
Sequence: 1 gaattcgacagggccacc.....agcggctcgtcggaattc 1932

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1932	100.0	1932	4	US-09-518-386B-2
2	353.2	18.3	3969	4	US-09-518-386B-4
3	48.4	2.5	1512	4	US-08-277-031B-4
4	48.4	2.5	2059	4	US-09-023-655-1062
5	48.4	2.5	2759	4	US-09-144-367-1
6	47.4	2.5	2160	4	US-09-716-129-12
7	46.8	2.4	1712	4	US-09-148-545-106
8	46.8	2.4	1822	4	US-09-148-545-105
9	45	2.3	1831	3	US-08-948-564-15
10	44	2.3	3828	4	US-09-221-013A-7
11	43.6	2.3	352	4	US-09-144-367-8
12	42.8	2.2	1333	4	US-09-372-422A-9
13	42.8	2.2	2407	3	US-09-370-807-7
14	42.8	2.2	2407	4	US-09-921-259-7
15	42.6	2.2	246240	2	US-08-724-394A-20
16	42.6	2.2	246240	2	US-08-724-394A-21
17	42.6	2.2	246240	2	US-08-724-394A-22
18	42.4	2.2	396	4	US-09-640-173-19
19	42.4	2.2	396	4	US-09-713-550-19
20	42.2	2.2	47	2	US-08-778-494B-114
21	42.2	2.2	323	4	US-09-621-976-10374
22	41.6	2.2	2202	4	US-09-396-149-3
23	41.4	2.1	1126	4	US-09-389-956-7
24	41.4	2.1	1507	4	US-09-453-323-1
25	41.4	2.1	7218	1	US-08-232-463-14
26	41.2	2.1	1024	4	US-09-328-475C-107
27	40.8	2.1	44848	4	US-09-435-739-42

28	40.6	2.1	1883	4	US-09-419-679-13	Sequence 13, Appli
29	40.2	2.1	947	4	US-09-673-395A-8	Sequence 8, Appli
30	40.2	2.1	5152	4	US-10-204-708-48	Sequence 48, Appli
31	40	2.1	155	4	US-09-621-976-11009	Sequence 11009, A
32	40	2.1	1260	4	US-09-674-741-5	Sequence 5, Appli
33	40	2.1	2581	2	US-09-013-634-1	Sequence 1, Appli
34	40	2.1	2745	4	US-09-833-381-1410	Sequence 1410, Ap
35	39.8	2.1	1599	4	US-09-023-655-1405	Sequence 1405, Ap
36	39.8	2.1	2099	4	US-09-800-729-56	Sequence 56, Appli
37	39.6	2.0	214	4	US-09-621-976-9843	Sequence 9843, Ap
38	39.6	2.0	396	4	US-08-640-173-16	Sequence 16, Appli
39	39.6	2.0	396	4	US-09-713-550-16	Sequence 16, Appli
40	39.6	2.0	1001	4	US-09-671-317-84	Sequence 84, Appli
41	39.6	2.0	2114	4	US-09-130-491-7	Sequence 7, Appli
42	39.6	2.0	3289	4	US-09-904-615-11	Sequence 11, Appli
43	39.6	2.0	6200	4	US-09-439-923-1	Sequence 1, Appli
44	39.6	2.0	6200	4	US-09-711-202A-1	Sequence 1, Appli
45	39.6	2.0	6200	4	US-09-711-205A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-518-386B-2
; Sequence 2, Application US/09518386B
; Patent No. 6365386
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518.386B
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1932
; TYPE: DNA
; ORGANISM: Phaffia rhodozyma
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (33)..(1706)
; NAME/KEY: polyA site
; LOCATION: (1871)
; NAME/KEY: rRNA
; LOCATION: (14)..(1891)
US-09-518-386B-2

Query Match	100.0%	Score 1932;	DB 4;	Length 1932;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1932;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GAATTCGGCAGAGCCACCTACTTCTCCATATGTTTCATCTTGGTCTTGTCTCACAGGTG	60	
DB	1	GAATTCGGCAGAGCCACCTACTTCTCCATATGTTTCATCTTGGTCTTGTCTCACAGGTG	60	
QY	61	CTTTAGGCTGGCTGCTTTCTCATGGGCATCCATAGCGTCTTTCAGTCTTTTACCTCGCTC	120	
DB	61	CTTTAGGCTGGCTGCTTTCTCATGGGCATCCATAGCGTCTTTCAGTCTTTTACCTCGCTC	120	
QY	121	CGAGCGCATCTTCACTGTATAACCTTCAGGCCCGCATCATACCACTACTTTACAGGCA	180	
DB	121	CGAGCGCATCTTCACTGTATAACCTTCAGGCCCGCATCATACCACTACTTTACAGGCA	180	
QY	181	ATTTTITAGACATCTCTCAGCTCGTACAGTGAAGAGCATGCGAAGTACAGAGAAAAT	240	
DB	181	ATTTTITAGACATCTCTCAGCTCGTACAGTGAAGAGCATGCGAAGTACAGAGAAAAT	240	

241 ACAGAGCAGCCCTCGGTTTCTCGGATCGCTGGAGCACCCTGTTGAATCTCGACCGATC 300
 Db |
 241 ACAGAGCAGCCCTCGGTTTCTCGGATCGCTGGAGCACCCTGTTGAATCTCGACCGATC 300
 Qy |
 301 CGAAAGTCTTCAACCATGTGATGAAGAAGCTACGACTATCCGAAACCTGGTATGCGCG 360
 Db |
 301 CGAAAGTCTTCAACCATGTGATGAAGAAGCTACGACTATCCGAAACCTGGTATGCGCG 360
 Qy |
 361 CTCGAGTCTCAGAAATGCTACCGAGATGGTGTGTGTACGGCGGAGGTGAAGCTCATA 420
 Db |
 361 CTCGAGTCTCAGAAATGCTACCGAGATGGTGTGTGTGTACGGCGGAGGTGAAGCTCATA 420
 Qy |
 421 AGCGACATCGAAGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGGTCC 480
 Db |
 421 AGCGACATCGAAGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGGTCC 480
 Qy |
 481 CAATTTCTTAGAAAAGGTATGAATCTGTGCAAGAGATGAGGAGATGCGGCTGAGA 540
 Db |
 481 CAATTTCTTAGAAAAGGTATGAATCTGTGCAAGAGATGAGGAGATGCGGCTGAGA 540
 Qy |
 541 AGGATATGCGCGTGGAGAGTCCGCGGTGAAAAGAGGCAACAGACTCGAGACCGAAG 600
 Db |
 541 AGGATATGCGCGTGGAGAGTCCGCGGTGAAAAGAGGCAACAGACTCGAGACCGAAG 600
 Qy |
 601 GAGTCGATGTAAGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
 Db |
 601 GAGTCGATGTAAGAGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
 Qy |
 661 TTGACTATAGAGGAGCTCGCTCGAGAACAGCAATGAGCTCTATGTCGCTTTTGTG 720
 Db |
 661 TTGACTATAGAGGAGCTCGCTCGAGAACAGCAATGAGCTCTATGTCGCTTTTGTG 720
 Qy |
 721 GACTTACCGATGGGTTGCTCCTACCTTGGACTCGTTCAAGGCTATCATGTGGATTTG 780
 Db |
 721 GACTTACCGATGGGTTGCTCCTACCTTGGACTCGTTCAAGGCTATCATGTGGATTTG 780
 Qy |
 781 TACCTTACTTCGAACTATGAACGGAGATGATGATACCTTTGACTCAGAGATAGCAG 840
 Db |
 781 TACCTTACTTCGAACTATGAACGGAGATGATGATACCTTTGACTCAGAGATAGCAG 840
 Qy |
 841 TTTCCCGACGAGTTGGGATCGAGCTTATGGAGCAAAAGAGCGCGGCTGTTGGCTCAG 900
 Db |
 841 TTTCCCGACGAGTTGGGATCGAGCTTATGGAGCAAAAGAGCGCGGCTGTTGGCTCAG 900
 Qy |
 901 CTTCCGATCAGGCTGTTGATGAAGAAGATGTTCAAGGTCGGGATATCCTAAGTCTCCTAG 960
 Db |
 901 CTTCCGATCAGGCTGTTGATGAAGAAGATGTTCAAGGTCGGGATATCCTAAGTCTCCTAG 960
 Qy |
 961 TGAGAGCAAAATCGCCGCAACCTGCTGAATCTCAAAAGCTGTCGATGAGGAGGTAC 1020
 Db |
 961 TGAGAGCAAAATCGCCGCAACCTGCTGAATCTCAAAAGCTGTCGATGAGGAGGTAC 1020
 Qy |
 1021 TCGCTCAGATCAGTAACCTGTTATTTGCTGATATGAATCTTTCGACAGCTTTCGACAT 1080
 Db |
 1021 TCGCTCAGATCAGTAACCTGTTATTTGCTGATATGAATCTTTCGACAGCTTTCGACAT 1080
 Qy |
 1081 GGATGTTTCCCGACTCTCAGAGACAAAGCGGTTCAAGTAAACTTCGAGAGAAATTT 1140
 Db |
 1081 GGATGTTTCCCGACTCTCAGAGACAAAGCGGTTCAAGTAAACTTCGAGAGAAATTT 1140
 Qy |
 1141 GTCAGATCGACACGATATGCTACGCTAGACGAATTAATGCGTGGCTTATCTCGAAG 1200
 Db |
 1141 GTCAGATCGACACGATATGCTACGCTAGACGAATTAATGCGTGGCTTATCTCGAAG 1200
 Qy |
 1201 CGTTTGTAAAGAGTCTCTGCTAGACCTCTAGTCCGTATGCTAAACCGTGAATGCT 1260
 Db |
 1201 CGTTTGTAAAGAGTCTCTGCTAGACCTCTAGTCCGTATGCTAAACCGTGAATGCT 1260
 Qy |
 1261 TAAAGGATGAAGACTTCAATCCACTTGCAGGCTGTCATTTGTCGAGATGGTGGTCA 1320
 Db |
 1261 TAAAGGATGAAGACTTCAATCCACTTGCAGGCTGTCATTTGTCGAGATGGTGGTCA 1320

1321 TCAACGAGGTCGGATCAGAAAGGAACGATGGTCTATGCTTCCGTTGTTCAACATCAATC 1380
 Db |
 1321 TCAACGAGGTCGGATCAGAAAGGAACGATGGTCTATGCTTCCGTTGTTCAACATCAATC 1380
 Qy |
 1381 GTTCAAAGTTTCAATTTATGAGAAAGATGAGAAAGATTCAGACCGAGAGGTGGCTTGAAG 1440
 Db |
 1381 GTTCAAAGTTTCAATTTATGAGAAAGATGAGAAAGATTCAGACCGAGAGGTGGCTTGAAG 1440
 Qy |
 1441 ACCTAACAGAGACTCGCTCAACAGATTTGAAGCAACCTATGGAACACCGAGGCTTATCT 1500
 Db |
 1441 ACCTAACAGAGACTCGCTCAACAGATTTGAAGCAACCTATGGAACACCGAGGCTTATCT 1500
 Qy |
 1501 CTGACCCAGAGCTTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1560
 Db |
 1501 CTGACCCAGAGCTTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1560
 Qy |
 1561 TTGTCACTCTCCGTCGGGTCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACA 1620
 Db |
 1561 TTGTCACTCTCCGTCGGGTCAGTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACA 1620
 Qy |
 1621 TCACCTTGATCATTTCCCGTCCCTCGAATCGTTGGTAGAGAAAGGAGGTACAGATGC 1680
 Db |
 1621 TCACCTTGATCATTTCCCGTCCCTCGAATCGTTGGTAGAGAAAGGAGGTACAGATGC 1680
 Qy |
 1681 GTTTCAGAGTCAAGCGGTCGATGATGATTTCTTCTCATATGTTAAGAGAGTTCTATAT 1740
 Db |
 1681 GTTTCAGAGTCAAGCGGTCGATGATGATTTCTTCTCATATGTTAAGAGAGTTCTATAT 1740
 Qy |
 1741 CTGAGATGTTGACTAGGACAAATGCTTCTTTGATCGAATTTGTTCTCATACCCGGC 1800
 Db |
 1741 CTGAGATGTTGACTAGGACAAATGCTTCTTTGATCGAATTTGTTCTCATACCCGGC 1800
 Qy |
 1801 AGGCGTATGACTCTTACGTCGCTATCGCTCTGAGCTCTCTTCTTACCTATATAT 1860
 Db |
 1801 AGGCGTATGACTCTTACGTCGCTATCGCTCTGAGCTCTCTTCTTACCTATATAT 1860
 Qy |
 1861 TATTCATCCGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAA 1920
 Db |
 1861 TATTCATCCGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAA 1920
 Qy |
 1921 CGTCCGCAATTC 1932
 Db |
 1921 CGTCCGCAATTC 1932

RESULT 2
 US-09-518-386B-4
 ; Sequence 4, Application US/095183865
 ; Patent No. 6365386
 ; GENERAL INFORMATION:
 ; APPLICANT: HOSHINO, Tatsuo
 ; APPLICANT: OJIMA, Kazuyuki
 ; APPLICANT: SETOGUCHI, Yutaka
 ; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
 ; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
 ; CURRENT APPLICATION NUMBER: US/09/518,386B
 ; CURRENT FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: EP 99104668.1
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: EP 00101666.5
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 3969
 ; TYPE: DNA
 ; ORGANISM: Phaffia rhodozyma
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: (517)..(518)
 ; NAME/KEY: intron
 ; LOCATION: (784)..(898)
 ; NAME/KEY: intron

LOCATION: (1016)..(1087)
NAME/KEY: intron
LOCATION: (1180)..(1302)
NAME/KEY: intron
LOCATION: (1518)..(1600)
NAME/KEY: intron
LOCATION: (1635)..(1723)
NAME/KEY: intron
LOCATION: (1867)..(1939)
NAME/KEY: intron
LOCATION: (2000)..(2081)
NAME/KEY: intron
LOCATION: (2182)..(2257)
NAME/KEY: intron
LOCATION: (2355)..(2431)
NAME/KEY: intron
LOCATION: (2543)..(2618)
NAME/KEY: intron
LOCATION: (2653)..(2742)
NAME/KEY: intron
LOCATION: (2815)..(2962)
NAME/KEY: intron
LOCATION: (3051)..(3113)
NAME/KEY: intron
LOCATION: (3172)..(3247)
NAME/KEY: intron
LOCATION: (3322)..(3398)
NAME/KEY: intron
LOCATION: (3424)..(3513)
NAME/KEY: polyA site
LOCATION: (3865)..(3866)
NAME/KEY: intron
LOCATION: (653)..(734)
US-09-518-3868-4

Query Match 18.3%; Score 353.2; DB 4; Length 3969;
Best Local Similarity 99.2%; Pred. No. 2.1e-97;
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1514 TTGCTTTGGTGGCGATTGCTGTCGCGAGATGAAGCCCTTCTTTGTCACCTCCG 1573
DB 3508 TCGCAGTGTGGCGATTGCTGTCGCGAGATGAAGCCCTTCTTTGTCACCTCCG 3567

QY 1574 TCGGTCGAGTTCGAGCCCATCATCTCTCATCCAGATACGACACATCCTGATCAT 1633
DB 3568 TCGGTCGAGTTCGAGCCCATCATCTCTCATCCAGATACGACACATCCTGATCAT 3627

QY 1634 TTCCCGTCTCGAATCGTTGGTAGAGAGAGAGGGGTACAGATGCGTTTGCAGGTCAA 1693
DB 3628 TTCCCGTCTCGAATCGTTGGTAGAGAGAGAGGGGTACAGATGCGTTTGCAGGTCAA 3697

QY 1694 GCGGTCGAATGAGTTGATTCTTCATATGTTAAGAGAGTTCTATATCTGAGAAATGTGTG 1753
DB 3688 GCGGTCGAATGAGTTGATTCTTCATATGTTAAGAGAGTTCTATATCTGAGAAATGTGTG 3747

QY 1754 ACTAGACATGCTTCTTTGATGATGATTTCTCATACCGGGGAGGCGCTATGACT 1813
DB 3748 ACTAGACATGCTTCTTTGATGATGATTTCTCATACCGGGGAGGCGCTATGACT 3807

QY 1814 TCTACGTCGCTATCGCTGCTGACTCTCTTCTTACCTATATATTCCATCCG 1871
DB 3808 TCTACGTCGCTATCGCTGCTGACTCTCTTCTTACCTATATATTCCATCCG 3865

RESULT 3
US-08-277-031B-4
; Sequence 4, Application US/08277031B
; Patent No. 6620593
; GENERAL INFORMATION:
; APPLICANT: Hayashi, Koji
; APPLICANT: Sakaki, Toshiyuki
; APPLICANT: Yabasaki, Yoshiyasu
; APPLICANT: Komai, Koichiro

APPLICANT: Kaneko, Hideo
APPLICANT: Nakatsuka, Iwao
TITLE OF INVENTION: METHOD FOR SAFETY EVALUATION OF
TITLE OF INVENTION: CHEMICAL COMPOUND USING RECOMBINANT YEAST EXPRESSING
TITLE OF INVENTION: HUMAN CYTOCHROME P450
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5inch, 1.44MB
COMPUTER: IBM PC
OPERATING SYSTEM: Dos 5.0
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/277,031B
FILING DATE: 19-JULY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP-201120/1993
APPLICATION NUMBER: JP-180246/1993
APPLICATION NUMBER: JP-208279/1993
FILING DATE: 20-07-1993
FILING DATE: 21-07-1993
FILING DATE: 30-07-1993
ATTORNEY/AGENT INFORMATION:
NAME: Raymond C. Stewart
REGISTRATION NUMBER: 21,066
REFERENCE/DOCKET NUMBER: 20-3530P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-277-031B-4

Query Match 2.5%; Score 48.4; DB 4; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.00023;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATG 1056
DB 863 AAGCTCTGTCCGATGAGGCTCGTGGCCCAATCAATTAATTTATTTGCTGGCTATG 922

QY 1057 AAACTTCTTCGACAGTCTTTGACATGATGTTTCCCGACTCTCAGAAAGAACCCGTTTC 1116
DB 923 AAACACGAGCAGTGTCTCTCTCTTATGATGATGAACTGGCCACTCACCTGATGTC 982

QY 1117 AGGTAATACTTCGAGAGAAATTTGTGATGCGACAGGATATGCTTACGCTAGACGAAC 1176
DB 983 AGCAGAACTTCGAGGAGAAATTTGATGAGTTTACCCATTAAGGACCCACCTATG 1042

QY 1177 TTAATGCGGTTCG-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTGTCTGAGCC 1230
DB 1043 ATACTGTGCTACAGATGAGTATCTTGATGTTGGTGAATGAAACGCTCAGATTATTC 1102

QY 1231 CTCCTAGTCCGATGCTTAACGCTGAATGCTTAAAGGATGAGACTTCA 1278
DB 1103 CAATTGCTATGAGACTTGAGAGGGTCTCGAAAAAAGATGTTGAGATCA 1150

RESULT 4
US-09-023-655-1062
; Sequence 1062, Application US/09023655


```

; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 1712

```

```

Query Match      2.4%; Score 46.8; DB 4; Length 1712;
Best Local Similarity 85.0%; Pred. NO. 0.00077;
Matches 51; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY      1854 TATATATTATTCATCCGAAAAAAAAAAAAAAAAAGCGCGCTCGA 1913
          |||||
Db       1645 TAATATATTTTCTACAAAAAAAAAAAAAAAAAGSGCGCTCGA 1704

RESULT 8
US-09-148-545-105
; Sequence 105, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; EARLIER APPLICATION NUMBER: 1998-09-04
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23

```

EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 105
LENGTH: 1822

Query Match 2.4%; Score 46.8; DB 4; Length 1822;

Best Local Similarity 85.0%; Pred. No. 0.0008;

Matches 51; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1854 TATATATATTCATCCGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGA 1913

Db 1755 TAAATATATTTTGTACAAAAAAAAAAAAAAAAAAAAAAAAAGSGCGCTCGA 1814

```
RESULT 9
US-08-948-564-15
; Sequence 15, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminazky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; METHODS OF PRODUCING Herbicide-Resistant Transgenic Plants
; NUMBER OF INVENTIONS: 23
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..1747
; US-08-948-564-15
;
; Query Match 2.3%; Score 45; DB 3; Length 1831;
; Best Local Similarity 52.8%; Pred. No. 0.0029;
; Matches 121; Conservative 0; Mismatches 105; Indels 3; Gaps 1;
;
; QY 1008 GATGAGAGGACTCGCTCAGATCAGTACCTGTTATTGTCGTGATATGAACCTTCTTCG 1067
; Db |||||
; 1037 GATCGTCAGTTGAGGATGATTAATGACAATGCTTATTGCGGTCATGAAACACGCT 1096
; QY |||||
; 1068 ACAGTCTTGACATGGATGTTTACCGACTCTCAGAGACAAAGCGTTTCAGGATAAATT 1127
; Db |||||
; 1097 GCAGTTCTTACTTGGCAGCTTTTCTCCTAGCTCAAATCTAGCAAAATGAAGAGCT 1156
; QY |||||
; 1128 CGAGAGA---AATTGTCAGATCGACACGATATGCTTACGCTAGACGAACTTAATGCG 1184
; Db |||||
; 1157 CAAGCAGAGTAGATTGCTGCTGGGTACGGGAGGCCAACTTTTGAATCATCTTAAGGAA 1216
; QY |||||
; 1185 TTGCTTATCTCGAAGCGTTTGTAAAGGAGTCTTTCGTCTAGACCCCTC 1233
; Db |||||
; 1217 TTGCAGTACATTAAGATTGTTGAGGAGCTCTTCGTTTATACCCCTC 1265
;
; RESULT 10
US-09-221-013A-7
; Sequence 7, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
; APPLICANT: Arioli, Antonio
; APPLICANT: Williamson, Richard E.
; APPLICANT: Betzner, Andreas S.
; APPLICANT: Peng, Liangcai
; TITLE OF INVENTION: Manipulation of cellulose and/or Beta-1,4-glucan
; FILE REFERENCE: 96-98
; CURRENT APPLICATION NUMBER: US/09/221,013A
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: PCT/AU97/00402
; PRIOR FILING DATE: 1997-06-24
; PRIOR APPLICATION NUMBER: AU P00699
; PRIOR FILING DATE: 1996-06-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3828
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (239)..(3490)
; US-09-221-013A-7
;
; Query Match 2.3%; Score 44; DB 4; Length 3828;
; Best Local Similarity 83.3%; Pred. No. 0.0092;
; Matches 50; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
;
; QY 1854 TATATATTATTCATCCGAAAAAAGGCGCGCTCGA 1913
; Db |||||
; 3746 TACACATTTGTTTACAAAAAAGGCGCGCTCGA 3805
;
; RESULT 11
US-09-144-367-8
; Sequence 8, Application US/09144367
; Patent No. 6432639
; GENERAL INFORMATION:
; APPLICANT: Lichter, Jay
; APPLICANT: Guido, Marco
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-12P
; CURRENT APPLICATION NUMBER: US/09/144,367
; CURRENT FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/058,612
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 352
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (0)...(0)
; US-09-144-367-8
;
; Query Match 2.3%; Score 43.6; DB 4; Length 352;
; Best Local Similarity 57.2%; Pred. No. 0.0027;
; Matches 79; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
;
; QY 1002 CTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTTGTGATATGAAC 1061
; Db |||||
; 123 CTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTTGTGATATGAAC 182
; QY |||||
; 1062 TCTTCGACAGTCTTGACATGGATGTTTACCGACTCTCAGAGACAAAGCGTTTCAGAT 1121
; Db |||||
; 183 ACAGCAGTGTCTCTCTCTTATGTAATGCACTGCCACTCACCTGATGTCAGAG 242
; QY |||||
; 1122 AAACCTCGAAGAAATT 1139
; Db |||||
; 243 AAACCTCGAAGAAATT 260
```



```

; Patent No. 6465234
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/921,259
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/096,225
; PRIOR FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-921-259-7

Query Match      2.2%; Score 42.8; DB 4; Length 2407;
Best Local Similarity 80.6%; Pred. No. 0.016;
Matches 50; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1840 CTCCTCTTACCTATATATTTCCATCCGAAAAA 1999
Db 1939 CTATCATCTTAGCGTTAATCTCTGCGATGAAAAA 1998
QY 1900 AA 1901
Db 1999 AA 2000

RESULT 15
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872337
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolffe, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872337el
; TITLE OF INVENTION: Sequences and Antibodies Thereof
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:

; Patent No. 6465234
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/921,259
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/096,225
; PRIOR FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-921-259-7

Query Match      2.2%; Score 42.8; DB 4; Length 1333;
Best Local Similarity 87.0%; Pred. No. 0.011;
Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1855 ATATATTTTCATCCGAAAAA 1998
Db 1280 ATCGATTATTTCCCGCAAAAAA 1333

RESULT 13
US-09-370-807-7
; Sequence 7, Application US/09370807
; Patent No. 6237034
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/370,807
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/096,225
; EARLIER FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-370-807-7

Query Match      2.2%; Score 42.8; DB 3; Length 2407;
Best Local Similarity 80.6%; Pred. No. 0.016;
Matches 50; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1840 CTCCTCTTACCTATATATTTCCATCCGAAAAA 1999
Db 1939 CTATCATCTTAGCGTTAATCTCTGCGATGAAAAA 1998
QY 1900 AA 1901
Db 1999 AA 2000

RESULT 14
US-09-921-259-7
; Sequence 7, Application US/09921259
```

```
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match      2.2%; Score 42.6; DB 2; Length 246240;
Best Local Similarity 67.4%; Pred. No. 0.34;
Matches 50; Conservative Indels 0; Gaps 0;

QY 1813 TTCTAGTCGTCTATCGTCTGGCTCTCTTACCCCTATATATTTCCATCCGA 1872
DB 185758 TCCTACTGCCCATAAAGACAGGCAAACTCCTGGGCTTCCCTGTATATGCTTCCCGCAA 185817

QY 1873 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1901
DB 185818 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 185846
```

Search completed: March 31, 2004, 10:41:39
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 13:24:50 ; Search time 742 Seconds

(without alignments)
9758.296 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932
Sequence: 1 gaattcgccagggccacc.....agccggctctgcgaattc 1932

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2470632 seqs, 187387510 residues

Total number of hits satisfying chosen parameters: 4941264

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1932	100.0	1932	14	US-10-066-007-2
2	353.2	18.3	3969	14	Sequence 2, Appli
3	49.8	2.6	467	12	Sequence 4, Appli
4	49.4	2.6	1501	15	Sequence 20949, A
5	49.2	2.5	368	12	Sequence 392, App
6	48.4	2.5	657	12	Sequence 1793, Ap
7	48.4	2.5	1512	15	Sequence 25496, A
8	48.4	2.5	2011	9	Sequence 55, Appl
9	48.4	2.5	2759	14	Sequence 1586, Ap
10	48.4	2.5	2768	14	Sequence 1, Appli
11	48.4	2.5	2849	9	Sequence 14, Appl
12	48.2	2.5	312	9	Sequence 2110, Ap
13	47.6	2.5	3059	9	Sequence 4611, Ap
14	47.4	2.5	628	12	Sequence 268, App
15	47.4	2.5	1410	12	Sequence 71313, A
					Sequence 29, Appl

c 16	47.2	2.4	1228	12	US-10-424-599-32438
c 17	47	2.4	218	9	US-09-933-797-162
c 18	46.8	2.4	685	12	US-10-424-599-8656
c 19	46.8	2.4	709	12	US-10-424-599-75080
20	46.8	2.4	1151	12	US-10-424-599-23175
21	46.8	2.4	1712	9	US-09-981-876-106
22	46.8	2.4	1712	10	US-09-148-545-106
23	46.8	2.4	1822	9	US-09-981-876-105
24	46.8	2.4	1822	10	US-09-148-545-105
c 25	46.6	2.4	1091	12	US-10-424-599-72966
c 26	46.4	2.4	1419	12	US-10-424-599-106636
c 27	46.2	2.4	672	12	US-10-424-599-99744
c 28	46.2	2.4	710	12	US-10-424-599-15121
29	46.2	2.4	1194	12	US-10-424-599-29930
c 30	46	2.4	266	12	US-10-424-599-64827
31	46	2.4	414	12	US-10-424-599-94220
32	46	2.4	940	12	US-10-424-599-111262
c 33	45.8	2.4	580	12	US-10-066-543-950
c 34	45.8	2.4	592	14	US-10-066-543-950
c 35	45.8	2.4	991	12	US-10-424-599-54757
36	45.8	2.4	1498	12	US-10-424-599-54758
37	45.8	2.4	2301	12	US-10-424-599-138418
c 38	45.6	2.4	447	12	US-10-424-599-142657
c 39	45.6	2.4	583	12	US-10-424-599-52047
c 40	45.6	2.4	705	12	US-10-424-599-137273
41	45.6	2.4	733	12	US-10-424-599-118266
42	45.4	2.3	558	12	US-10-424-599-97712
43	45.4	2.3	662	12	US-10-424-599-67477
44	45.4	2.3	767	12	US-10-424-599-80087
45	45.4	2.3	1359	15	US-10-369-493-43674

ALIGNMENTS

RESULT 1

US-10-066-007-2
Sequence 2, Application US/10066007
Publication No. US20030077691A1

GENERAL INFORMATION:

APPLICANT: HOSHINO, Tatsuo
APPLICANT: OJIMA, Kazuyuki
APPLICANT: SETOGUCHI, Yutaka
TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
FILE REFERENCE: ASTAXANTHIN SYNTHETASE
CURRENT APPLICATION NUMBER: US/10/066,007
CURRENT FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: US/09/518,386
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: EP 99104668.1
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: EP 00101666.6
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1932
TYPE: DNA

ORGANISM: Phaffia rhodozyma

FEATURE:

NAME/KEY: CDS

LOCATION: (33)..(1706)

NAME/KEY: polyA site

LOCATION: (1871)

NAME/KEY: mRNA

LOCATION: (14)..(1891)

US-10-066-007-2

Query Match 100.0%; Score 1932; DB 14; Length 1932;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGCAGGAGCCACTTCTCCATATGTTTCATCTTGCTGCACAGTG 60

1 GAAATCGGCACGAGCCACCTACTTTCTCCATATGTTTCATCTTGGCTCTGCACAGGTG 60
61 CTTTAGGCTCGGCTGCTTTCTCATGGGCTCCATAGGTTCTTCACTCTTACCTCGCTC 120
61 CTTTAGGCTCGGCTGCTTTCTCATGGGCTCCATAGGTTCTTCACTCTTACCTCGCTC 120
121 CGAGCGATCTTCACTGATTAACCTTCAGGCGCGGAATCATACAACTACTTTACAGGCA 180
121 CGAGCGATCTTCACTGATTAACCTTCAGGCGCGGAATCATACAACTACTTTACAGGCA 180
181 ATTTTATAGCATCTCTCACTCGTACAGTGAAGAGCTGGAAGTACAGAGAAAAT 240
181 ATTTTATAGCATCTCTCACTCGTACAGTGAAGAGCTGGAAGTACAGAGAAAAT 240
241 ACGGAAGCACCTCCGGTTTCTCGGATTCGCTGAGCACCCGCTTTGAATCGACCGATC 300
241 ACGGAAGCACCTCCGGTTTCTCGGATTCGCTGAGCACCCGCTTTGAATCGACCGATC 300
301 CGAAAGTCTTCAACCATGATGAAGAGCTTACGACTATCGAAACCTGGTATGCGCG 360
301 CGAAAGTCTTCAACCATGATGAAGAGCTTACGACTATCGAAACCTGGTATGCGCG 360
361 CTCGAGTGTCTCAGAAATGCTTACCGGAGATGCTGTGTACGCGGGAAGGTGAAGCTCAT 420
361 CTCGAGTGTCTCAGAAATGCTTACCGGAGATGCTGTGTGTACGCGGGAAGGTGAAGCTCAT 420
421 ACGGACATCGAAGGATCATGATCCCTCTCTGTCGCTCAGGCGGTTAACTCGATGGTCC 480
421 ACGGACATCGAAGGATCATGATCCCTCTCTGTCGCTCAGGCGGTTAACTCGATGGTCC 480
481 CAATTTCTTGAAGAGTATGGAACCTGTGCAAGAGATGATGGAGGATGCGGCTGAGA 540
481 CAATTTCTTGAAGAGTATGGAACCTGTGCAAGAGATGATGGAGGATGCGGCTGAGA 540
541 AGGATATGCGGCTGGGAGAGTGGGCTCGAGTACTCTGAGCTCATGGCTCTTGAGGAT 600
541 AGGATATGCGGCTGGGAGAGTGGGCTGGAGAGGGAACAGACTCGAGACCGAAG 600
601 GAGTCGATTAAGGATTTGGTGGCTCGAGTACTCTGAGCTCATGGCTCTTGAGGAT 660
601 GAGTCGATTAAGGATTTGGTGGCTCGAGTACTCTGAGCTCATGGCTCTTGAGGAT 660
661 TTGACTATTAAGAGCGACTCGCTCCAGAACAGACCAATGAGCTCTATGTGCTTTTTCG 720
661 TTGACTATTAAGAGCGACTCGCTCCAGAACAGACCAATGAGCTCTATGTGCTTTTTCG 720
721 GACTTACCGATGGGTTTGTCTCTTACCTTGGACTGCTCAAGGCTCATGTGGGATTTG 780
721 GACTTACCGATGGGTTTGTCTCTTACCTTGGACTGCTCAAGGCTCATGTGGGATTTG 780
781 TACCTTTACTTCCGAATATGAACGAGACATGAGATACCTTTGACTCAAGGATTAGCAG 840
781 TACCTTTACTTCCGAATATGAACGAGACATGAGATACCTTTGACTCAAGGATTAGCAG 840
841 TTTCCGAGAGTGGATGGATTAAGGACCAAGAGCGGCTGCTGGCTCAG 900
841 TTTCCGAGAGTGGATGGATTAAGGACCAAGAGCGGCTGCTGGCTCAG 900
901 CTTCCGATCAGGCTGTGTATATAAAGGATGTTCAAGGTCGGGATATCTTAAGTCTCCTAG 960
901 CTTCCGATCAGGCTGTGTATATAAAGGATGTTCAAGGTCGGGATATCTTAAGTCTCCTAG 960
961 TGAGAGCAACATCGCGGCAACCTGCTGATCTCAAAAGCTGTCGATGAGGATGAC 1020
961 TGAGAGCAACATCGCGGCAACCTGCTGATCTCAAAAGCTGTCGATGAGGATGAC 1020
1021 TCGCTCAGATCAGTAACCTGTTATTTGCTGGATATGAACCTCTTCGACAGCTTTGACAT 1080
1021 TCGCTCAGATCAGTAACCTGTTATTTGCTGGATATGAACCTCTTCGACAGCTTTGACAT 1080
1081 GGATGTTTCAACGACTCTCAGAGAACAAAGCGGTCAGGATAACTTCGAGAGAAATTT 1140

1081 GGATGTTTCAACGACTCTCAGAGAACAAAGCGGTTCCAGATAAACTTCGAGAGAAATTT 1140
1141 GTGAGATCGACACGGATATGCTTACGCTAGAGAACTTAAATGCGTTGCTTATCTCGAAG 1200
1141 GTGAGATCGACACGGATATGCTTACGCTAGAGAACTTAAATGCGTTGCTTATCTCGAAG 1200
1201 CGTTTGTAAAGGAGTCTCTTCTGCTAGACCTCTTAGTCCGATATCTAACCGTGAATGCT 1260
1201 CGTTTGTAAAGGAGTCTCTTCTGCTAGACCTCTTAGTCCGATATCTAACCGTGAATGCT 1260
1261 TAAAGATGAAGACTTTCATCCCATCTTCCGAGCTTCAATGCTCGAGATGGTGGTCA 1320
1261 TAAAGATGAAGACTTTCATCCCATCTTCCGAGCTTCAATGCTCGAGATGGTGGTCA 1320
1321 TCAACGAGTCCGAGTACGAGAGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
1321 TCAACGAGTCCGAGTACGAGAGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
1381 GTTCAAAGTTCAATTTATGGAGAAAGATGCAAGAAATTCAGACCGAGAGGTGGTGGG 1440
1381 GTTCAAAGTTCAATTTATGGAGAAAGATGCAAGAAATTCAGACCGAGAGGTGGTGGG 1440
1441 ACGTAAAGAGTCCGCTCAACAGTATTGAAGCACTTATGACACCAAGGCGAGCTTATCT 1500
1441 ACGTAAAGAGTCCGCTCAACAGTATTGAAGCACTTATGACACCAAGGCGAGCTTATCT 1500
1501 CTGACCCAGAGTGGTGGTGGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
1501 CTGACCCAGAGTGGTGGTGGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
1561 TTGTCACCTCTCGGCTCGGCTCGAGTTCGAGCCCATCATCTCATCCAGAGTACGAGCACA 1620
1561 TTGTCACCTCTCGGCTCGGCTCGAGTTCGAGCCCATCATCTCATCCAGAGTACGAGCACA 1620
1621 TCACCTTGATCAATTTCCGCTCGAATCGTGGTAGAGAGAGGAGGCTACAGATGC 1680
1621 TCACCTTGATCAATTTCCGCTCGAATCGTGGTAGAGAGAGGAGGCTACAGATGC 1680
1681 GTTTCGAGGTCAGCCGCTCGAATGAGTTCATATCTTATGATGTAAGAGAGTCTTATAT 1740
1681 GTTTCGAGGTCAGCCGCTCGAATGAGTTCATATCTTATGATGTAAGAGAGTCTTATAT 1740
1741 CTGAGATGCTGACTAGGACAAATGCTTCTTGTATCGAATTTGTTCTCATACCGGCGC 1800
1741 CTGAGATGCTGACTAGGACAAATGCTTCTTGTATCGAATTTGTTCTCATACCGGCGC 1800
1801 AGGCGCTATGACTTCTACGCTGCTATCGTGGCTCTGCTGCTCTCTCTTCTTACCTATAT 1860
1801 AGGCGCTATGACTTCTACGCTGCTATCGTGGCTCTGCTGCTCTCTCTTCTTACCTATAT 1860
1861 TATTCCTCCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
1861 TATTCCTCCGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
1921 CGTCCGGAATTC 1932
1921 CGTCCGGAATTC 1932

RESULT 2

US-10-066-007-4
; Sequence 4, Application US/10066007
; Publication No. US20030077691A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yukaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/10/066,007
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US/09/518,386
; PRIOR FILING DATE: 2000-03-03

QY 1845 TTCTTACCTATATATATCCATCCGAGAAAAA 1901
Db 1430 TTTTCATATAAGATTGATTTTAAACAAAAA 1486

RESULT 5
US-10-424-599-1793
; Sequence 1793, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 1793
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101618C.1
US-10-424-599-1793

Query Match 2.5%; Score 49.2; DB 12; Length 368;
Best Local Similarity 87.1%; Pred. No. 0.00024;
Matches 54; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1853 CTATATATATTTCCATCCGAAAAA 1912
Db 299 CAATGATACCTTACCGAAAAA 358
QY 1913 AG 1914
Db 359 AG 360

RESULT 6
US-10-424-599-25496/c
; Sequence 25496, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 25496
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_123024C.1
US-10-424-599-25496

Query Match 2.5%; Score 48.4; DB 12; Length 657;
Best Local Similarity 74.4%; Pred. No. 0.00061;
Matches 61; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1833 CTCTGACCTCTCTTACCTATATATTTCCATCCGAAAAA 1892
Db 91 CTCTGAGATCTTTCGAAACGATTCAGATGAAGACAAAAA 32
QY 1893 AAAAAAAGCGCGCTCGAG 1914

Db 31 AAAAAAAGCGCGCTCTAG 10

RESULT 7
US-10-313-963A-55
; Sequence 55, Application US/10313963A
; Publication No. US20040002078A1
; GENERAL INFORMATION:
; APPLICANT: Boutell, Jonathan
; APPLICANT: Godber, Benjamin
; APPLICANT: Hart, Darren
; APPLICANT: Blackburn, Jonathan
; TITLE OF INVENTION: Arrays
; FILE REFERENCE: KIL-001
; CURRENT APPLICATION NUMBER: US/10/313,963A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/335,806
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/410,815
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 1512
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-313-963A-55

Query Match 2.5%; Score 48.4; DB 15; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.001;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
QY 997 AAAAGCTGTCGATGAGGAGTACTCGCTCAGATCACTAGTAACCTGTTATTTGCTGATG 1056
Db 863 AAGCTCTGTCGATCTGAGCTCGTGGCCCAATCAATATCTTTATTTTGTGGCTATG 922
QY 1057 AAATTTCTTCGACAGCTTTGACATGGATGTTTCACGACCTCTCAGAGACAAGCCGTTTC 1116
Db 923 AAACGACGAGCAGTTCTCTCTCTCATTTATGATGAATGCGCACTACCTCGATGTC 982
QY 1117 AGGATAAAGCTTCGAGAGAAATTTGTCAGATCGACACGATATGCTAGCTAGACGAAAC 1176
Db 983 AGCAGAACTGCGAGGAGAAATTTGATGTCAGTTTACCAATAGGACCAACCCCTATG 1042
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTCTGCTAGACC 1230
Db 1043 ATACTGTCTACAGATGGAGTATCTTGACATGCTGTGTAATGAACGCTCAGATTATTC 1102
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1103 CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAAGATGTTGATCA 1150

RESULT 8
US-09-880-107-1586
; Sequence 1586, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 1586
; LENGTH: 2011
; TYPE: DNA
; ORGANISM: Homo sapiens
; PEATRE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00003
US-09-880-107-1586

Query Match
Best Local Similarity 50.7%; Pred. No. 0.0012;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAAACCTGTTATTTGCTGGATATG 1056
Db 926 AAGCTGTGCGATCTGGAGTCTGGGCCAATCAATATCTTTATTTTCTGGCTATG 985
QY 1057 AAACCTTCTTCGACAGCTTTGATGATGTTTCCACGAGTCTCAGAAAGAAAGCGGTTT 1116
Db 986 AAACCAACGAGCAGTGTCTCTCTTCATATGATGATGAACTGGCCACTCAACCTGATGTC 1045
QY 1117 AGATTAACCTTCGAGAAATTTGTCAGATCGACACGGATATGCTTACCTAGACGAAC 1176
Db 1046 AGCAGAACTGCGAGGAAATTTGATGAGTTTACCAATAGGCACACCCACCTATG 1105
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGAAC 1230
Db 1106 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1165
QY 1231 CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db 1166 CAATTGCTATGAGATTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1213

```

```

RESULT 9
US-10-146-575-1
; Sequence 1, Application US/10146575
; Publication No. US20030059800A1
; GENERAL INFORMATION:
; APPLICANT: Lichter, Jay
; APPLICANT: Guido, Marco
; TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4
; FILE REFERENCE: SEQ-12P
; CURRENT APPLICATION NUMBER: US/10/146,575
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: US/09/144,367
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2759
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)....(1581)
; OTHER INFORMATION: Human CYP3A4 cDNA reference sequence
US-10-146-575-1

```

```

Query Match
Best Local Similarity 50.7%; Pred. No. 0.0015;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAAACCTGTTATTTGCTGGATATG 1056
Db 932 AAGCTGTGCGATCTGGAGTCTGGGCCAATCAATATCTTTATTTTCTGGCTATG 991
QY 1057 AAACCTTCTTCGACAGCTTTGATGATGTTTCCACGAGTCTCAGAAAGAAAGCGGTTT 1116
Db 992 AAACCAACGAGCAGTGTCTCTCTTCATATGATGATGAACTGGCCACTCAACCTGATGTC 1051
QY 1117 AGATTAACCTTCGAGAAATTTGTCAGATCGACACGGATATGCTTACCTAGACGAAC 1176
Db 1052 AGCAGAACTGCAGGAGGAAATTTGATGAGTTTACCCAAATAGGACCAACCCCTATG 1111

```

```

QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGAAC 1230
Db 1112 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1171
QY 1231 CTCCTAGTCCGATGCTTAAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1172 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1219

RESULT 10
US-10-268-822-14
; Sequence 14, Application US/10268822
; Publication No. US20030150004A1
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Wei, Ping
; APPLICANT: Chua, Steven
; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob
; TITLE OF INVENTION: Metabolism
; FILE REFERENCE: P02729US2
; CURRENT APPLICATION NUMBER: US/10/268,822
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US 01/29672
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 10/219,590
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 2768
; TYPE: DNA
; ORGANISM: Human
US-10-268-822-14

```

```

Query Match
Best Local Similarity 50.7%; Pred. No. 0.0015;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAAACCTGTTATTTGCTGGATATG 1056
Db 967 AAGCTGTGCGATCTGGAGTCTGGGCCAATCAATATCTTTATTTTCTGGCTATG 1026
QY 1057 AAACCTTCTTCGACAGTCTTGACATGGATGTTTACCGACTCTCAGAAAGAAAGCGGTTT 1116
Db 1027 AAACCAACGAGCAGTGTCTCTCTTCATATGATGATGAACTGGCCACTCAACCTGATGTC 1086
QY 1117 AGGATAAACTTCGAGAAATTTGTCAGATCGACACGGATATGCTTACCTAGACGAAC 1176
Db 1087 AGCAGAACTGCAGGAGGAAATTTGATGAGTTTACCCAAATAGGACCAACCCCTATG 1146
QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGAAC 1230
Db 1147 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1206
QY 1231 CTCCTAGTCCGATGCTTAAACCGTGAATGCTTAAAGGATGAAGACTTCA 1278
Db 1207 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1254

```

```

RESULT 11
US-09-880-107-2110
; Sequence 2110, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14

```



```

, PRIOR APPLICATION NUMBER: US 60/211,379
,
, PRIOR FILING DATE: 2000-06-14
,
, PRIOR APPLICATION NUMBER: US 60/237,054
,
, PRIOR FILING DATE: 2000-10-02
,
, NUMBER OF SEQ ID NOS: 3350
,
, SOFTWARE: Patent In Ver. 2.1
,
, SEQ ID NO 1110
,
, LENGTH: 2849
,
, TYPE: DNA
,
, ORGANISM: Homo sapiens
,
, FEATURE:
,
, OTHER INFORMATION: Genbank Accession N
US-09-880-107-2110

```

Query Match	2.5%	Score	48.4	DB	9	Length	2849
Best Local Similarity	50.7%	Pred. No.	0.0015				
Matches	146	Conservative	0	Mismatches	136	Indels	6
Gaps	1						
Qy	997	AAAAGCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG	1056				
Db	928	AAAGCTCTGTCGATCGGAGCTCGTGGCCCAATATATCTTTATTTTGGTGGCTATG	987				
Qy	1057	AAACTTCTTCGACGCTCTTGACATGATGTTTCAACGACCTCTCAGAACGACAAAGCCGCTC	1116				
Db	988	AAACCAACGAGCAGTGTCTCTCTCTTCATATGATGAACTGGCCCACTCACCTGATGTCC	1047				
Qy	1117	AGGATAAACTTCGAGAAAGAAATTTGTCAGATCGACACGGATATGCTCAGCTAGACGAAAC	1176				
Db	1048	AGCAGAAACTGCAGGAGGAAATTCATGCAGTCTTTTACCAATAAGGCACCAACCCTATG	1107				
Qy	1177	TTAATGCGTTC-----CTTATCTCGAAGCGTTGTTAAGGAGTCTCTTGCTCTAGACC	1230				
Db	1108	ATACTGTGCTACAGATGGAGTATCTTGACATGTTGGTGAATGAACGCTCAGATTATCC	1167				
Qy	1231	CTCCTAGTCGCTATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA	1278				
Db	1168	CAATTGCTATGAGCTCAGAGGCTCTGCAAAAAGATGTTGAGATCA	1215				

RESULT 12
US-09-960-352-4611/c
; Sequence 4611, Application US/09960352
; Patent NO. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tac, Mengbing
; APPLICANT: Bysett, John C.
; APPLICANT: Machidagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4611
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3058-022-Q1-K1-E7
US-09-960-352-4611

	Query Match	2.5%;	Score 48.2;	DB 9;	Length 312;
	Best Local Similarity	55.8%;	Pred. No. 0.00044;		
	Matches 92;	Conservative	0;	Mismatches 73;	Indels 0; Gaps 0;
QY	1745	GAATGTTGTGACTAGGACAATGCTTTCTTTGTATCGATTTGTTTCTCATACCGGCGCAGGC	1804		
DB	175	GAATGCGAGGAGGAAAACCTGAGACCTTAGAGTATTAAATGTAGTGTCAAACCTAGCCATA	116		
QY	1805	GCTATGACACTTTCACGTCGCTATCGTCGCTGACACTCTCTTCTTACCCCTATATATTATT	1864		
DB	115	TGATAGACTGATATCATGTATGATCTGATGATCTTTTCTGCTTTAATAAAAAACCATACGATT	56		

Qy 1865 CCATCCGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGGCGC 1909

Dh 55 ACATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAGGACGCGC 11

```

RECORD 13
US-09-925-297-268
; Sequence 268, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 268
; LENGTH: 3059
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (432)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-268

```

	Query Match	2.5%	Score 47.6;	DB 9;	Length 3059;
	Best Local Similarity	92.6%;	Pred. No. 0.0029;		
	Matches 50;	Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1860	TTATTCCATCCG	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAGCGCGCTCGA	1913
hb	2998	TTATTCCTCCCA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAGCGCGCTCGA	3051

```

RESULT 14
US-10-424-599-71313/c
; Sequence 71313, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; AND/OR USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 71313
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35409C.1
US-10-424-599-71313

```

```

Query Match      2.5%; Score 47.4; DB 12; Length 628;
Best Local Similarity 89.5%; Pred. No. 0.0012;
Matches 51; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1858 TATTATTCATCCGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAG 1914
      |||||
Db       51 TATTAGTCAGCGGAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCTAG 5

RESULT 15

```

```

US-10-395-463-29
; Sequence 29, Application US/10395463
; Publication No. US20040060079A1
; GENERAL INFORMATION:
; APPLICANT: Tanaka, Hiroshi
; APPLICANT: Kayano, Toshiaki
; APPLICANT: Matsuoka, Makoto
; APPLICANT: Sakamoto, Tomoaki
; APPLICANT: Iwahori, Shuichi
; TITLE OF INVENTION: Method of Controlling Character of Monocotyledon by Modification
; TITLE OF INVENTION: Overexpression of Cytochrome P450 Monooxygenase Gene Involved in
; TITLE OF INVENTION: Biosynthesis and Monocotyledon Modified by the Gene
; FILE REFERENCE: 59150-8022
; CURRENT APPLICATION NUMBER: US/10/395,463
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-276398
; PRIOR FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1410)
; OTHER INFORMATION:
US-10-395-463-29

Query Match
Best Local Similarity 54.2%; Pred.No. 0.002;
Matches 96; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 997 AAAAGCTGTCGAGTGGAGGAGTCTCGCTCAGATCAGTAACTGTATTGCTGGATATG 1056
Db 791 AGAGCTCAGAGTGGAGGAGTCTCGCTCAGATCAGTAACTGTATTGCTGGATATG 850

QY 1057 AAACTTCTTCGACAGTCTTGACATGGATGTTTACCGACTCTCAGAGACAAAGCCGTC 1116
Db 851 AAACCATGTCGAGGACCTCGATGGTGTCAAGTACCTGTGACACCATCCCAAGCTC 910

QY 1117 AGGATAAATTCGAGAAAGAAATTTGTCAGATCGACAGGATATGCTACGCTAGACG 1173
Db 911 TTGAGCAACTCAGGAAGAACATTTTGATATCAGGAAGGTAAGCCCGAAGATG 967

```

Search completed: April 6, 2004, 17:58:23
Job time : 748 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 13:16:26 ; Search time 5193 Seconds
(without alignments)
11131.340 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932

Sequence: 1 gaattcgacagagccacc.....agccggctcgcggaattc 1932

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vit:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	4.6	528	14	CD275123 T143B0028
2	79.2	4.1	474	14	CD275883 T143B0523
3	79.2	4.1	475	14	CD275923 T143B0022
4	75.8	3.9	615	9	AT002896 AT002896

5	69.2	3.6	354	14	CD274181
6	68	3.5	403	14	CD273194
7	61	3.2	487	13	BQ102580
8	59.8	3.0	400	14	CD275427
9	58	3.0	317	14	CD275197
10	56.8	2.9	499	12	BJ075377
11	55.4	2.9	971	13	EX779027
12	53.8	2.8	849	14	CF152584
13	51.2	2.7	568	10	AW600900
14	50.6	2.6	1298	10	AW600865
15	50.4	2.6	875	14	CB200527
16	50.2	2.6	407	13	BU398544
17	50.2	2.6	723	13	BU250947
18	50.2	2.6	951	13	BU246013
19	50	2.6	740	12	BG207705
20	50	2.6	907	29	CNS015AK
21	49.8	2.6	212	9	AL651909
22	49.8	2.6	303	9	AA907131
23	49.6	2.6	544	14	CF386878
24	49.4	2.6	447	9	AV736322
25	49.2	2.5	387	13	BU779630
26	49	2.5	483	14	CB821156
27	48.8	2.5	455	9	AV744471
28	48.6	2.5	547	12	BI378359
29	48.4	2.5	396	12	BM139826
30	48.4	2.5	480	14	CB157627
31	48.4	2.5	580	14	CB114922
32	48.4	2.5	609	14	CB162338
33	48.4	2.5	641	14	CB154142
34	48.4	2.5	697	9	AV648984
35	48.4	2.5	755	12	BG206682
36	48.4	2.5	757	12	EG211348
37	48.4	2.5	760	12	EG196339
38	48.4	2.5	763	12	EG206681
39	48.4	2.5	765	12	EG182832
40	48.4	2.5	768	12	EG220280
41	48.4	2.5	770	12	EG197817
42	48.4	2.5	770	12	EG214545
43	48.4	2.5	789	12	EG188595
44	48.4	2.5	790	12	EG206683
45	48.4	2.5	791	12	EG208228

ALIGNMENTS

RESULT 1	CD275123	528 bp	mrna	linear	EST 01-SEP-2003
LOCUS	T143B00283 (FHIG:B)	Axenic plate culture	Paxillus involutus	CDNA	
DEFINITION	5', mRNA sequence.				
ACCESSION	CD275123				
VERSION	CD275123.1				
KEYWORDS	EST.				
SOURCE	Paxillus involutus				
ORGANISM	Paxillus involutus				
REFERENCE	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Boletales; Paxillaceae; Paxillus.				
AUTHORS	1 (bases 1 to 528)				
TITLE	Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R., Lundeborg, J., Uhlen, M. and Tunlid, A.				
JOURNAL	Transcriptional responses of Paxillus involutus and Betula pendula during formation of ectomycorrhizal root tissue				
COMMENT	Unpublished (2003)				
	Contact: Johansson, T.				
	Fungal-Host Interaction Group (FHIG)				
	Microbial Ecology, Institution of Ecology				
	Ecology Building, Lund University, SE-223 62 Lund, Sweden				
	Tel: +46 46 222 41 58				
	Fax: +46 46 222 41 58				
	Email: tomas.johansson@bioekol.lu.se				
	PCR Primers				
	FORWARD: P104 (5'-GGGAGCGCGCCATTGTGT-3')				

BACKWARD: P105 (5'-AGTGAGCTCGAATTGGGCC-3')

Seq primer: P104

High quality sequence stop: 528.

Location/Qualifiers

FEATURES

source

1. 528

/organism="Paxillus involutus"

/mol_type="rRNA"

/strain="ATCC 200175"

/db_xref="taxon:71150"

/tissue_type="Mycelium"

/dev_stage="25 days of growth after transfer"

/lab_host="Escherichia coli BM25.8"

/clone_lib="PHIG:B) Axenic plate culture"

/note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This

EST clone is originating from one of three cDNA libraries,

constructed for transcript profiling of the mycorrhizal

interaction between the basidiomycete *Paxillus involutus*

and *Betula pendula* (birch). One library represents the

developed and functional mycorrhizal root tissue

(PHIG:A) Ectomycorrhiza plate culture', a second

library represents axenically grown fungus (PHIG:B)

Axenic plate culture') and a third library represents

axenically grown plants (PHIG:C) Axenic plate culture'.

Libraries were analyzed in parallel and 3555 (PHIG:A),

3964 (PHIG:B), and 2532 (PHIG:C) high-quality (PHRED 20)

ESTs of >99bp have been deposited. The cDNA libraries were

constructed from total RNA using the SMART cDNA library

construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)

according to the manufacturer's instructions. Full-length

cDNAs were trimmed by SfiI, fractionated and directionally

ligated into (lambda)Triplex2 arms. The lambda library was

converted to a plasmid library via site-specific

recombination at loxP sites in a Cre+ strain (E. coli

BM25.8). Plasmid clones were randomly collected and

analysed by DNA sequencing using a plasmid-specific

forward primer (P104)."

ORIGIN

Query Match 4.6%; Score 88; DB 14; Length 528;

Best Local Similarity 54.5%; Pred. No. 8.5e-05;

Matches 244; Conservative 0; Mismatches 195; Indels 9; Gaps 3;

```

QY 1069 CAGCTTTGACATGGATGTTTCCACGACTCTCAGACGACAAAGCCGTTCCAGATAAATTC 1128
DB 1069 CAGCTTTGACATGGATGTTTCCACGACTCTCAGACGACAAAGCCGTTCCAGATAAATTC 1128
QY 1129 GAGAGAAATTTCTCAGATCGACGATATCGCTACGCTAGACGAACTTAATCGGTTGC 1188
DB 1129 GAGAGAAATTTCTCAGATCGACGATATCGCTACGCTAGACGAACTTAATCGGTTGC 1188
QY 1189 CTTATCTCGAAGGTTTGTAAAGGAGTCTCTTCGTCTAGACCTCTAGTCCGTTATGCTA 1248
DB 1189 CTTATCTCGAAGGTTTGTAAAGGAGTCTCTTCGTCTAGACCTCTAGTCCGTTATGCTA 1248
QY 208 CTTACCTTGACACCGTGTGAAGGAGACTTTTCGTGTACACCTCCCTTTGAGAGACCA 267
DB 208 CTTACCTTGACACCGTGTGAAGGAGACTTTTCGTGTACACCTCCCTTTGAGAGACCA 267
QY 1249 ACCGTGAATGTTAAAGGATGAAGACTTCATCCACTTCCGAGCTCTGATGTCGAG 1308
DB 1249 ACCGTGAATGTTAAAGGATGAAGACTTCATCCACTTCCGAGCTCTGATGTCGAG 1308
QY 268 TAGCGG---TCGGATGAAGACGATATCTTCGCTCGATGAGGACCTACGAGCAAC 324
DB 268 TAGCGG---TCGGATGAAGACGATATCTTCGCTCGATGAGGACCTACGAGCAAC 324
QY 1309 ATGGGTCGGTATCAACGAGGTCGGATCAGCAAGGAAACGATGTCATGTTCCGTTGT 1368
DB 1309 ATGGGTCGGTATCAACGAGGTCGGATCAGCAAGGAAACGATGTCATGTTCCGTTGT 1368
QY 325 ATGGGTCGGTATCAACGAGGAAATAGGATCAGCAAGGAAACGAGTCTCTATCCGATCT 384
DB 325 ATGGGTCGGTATCAACGAGGAAATAGGATCAGCAAGGAAACGAGTCTCTATCCGATCT 384
QY 1369 TCAACATCATGCTTCAAGATTCATTATTGAGAGAGATGCAGAGAAATTCAGACGAGGA 1428
DB 1369 TCAACATCATGCTTCAAGATTCATTATTGAGAGAGATGCAGAGAAATTCAGACGAGGA 1428
QY 385 T---GATGAACCGGTGGAAGGGGTTGTGGGGCCCTGACGCCACAGGTTCAACCCGGAAC 441
DB 385 T---GATGAACCGGTGGAAGGGGTTGTGGGGCCCTGACGCCACAGGTTCAACCCGGAAC 441
QY 1429 GGTGGCTTGAGGACGTAACAGACTCGCTCAACAGATATTGAAGACCCCTATGACACCCAGG 1488
DB 1429 GGTGGCTTGAGGACGTAACAGACTCGCTCAACAGATATTGAAGACCCCTATGACACCCAGG 1488
QY 442 GCTGG---GACGAGCTACCTAAAGCTGTCTCCAAATATCCAGGTGTTTGGGGCCACATGC 498
DB 442 GCTGG---GACGAGCTACCTAAAGCTGTCTCCAAATATCCAGGTGTTTGGGGCCACATGC 498
QY 1489 CGAGCTTTATCTCTGACACCCAGGCTTG 1516
DB 1489 CGAGCTTTATCTCTGACACCCAGGCTTG 1516
QY 499 TAAAGTTTCTCTGGCGGGCCGAGCATG 526
DB 499 TAAAGTTTCTCTGGCGGGCCGAGCATG 526

```

RESULT 2

CD275983

LOCUS

DEFINITION

CD275983

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CD275983

T143B05239F (PHIG:B) Axenic plate culture *Paxillus involutus* cDNA

5', mRNA sequence.

CD275983

CD275983.1 GI:34387929

EST.

Paxillus involutus

Paxillus involutus

Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

Boletales; Paxillaceae; Paxillaceae; Paxillus.

1 (bases 1 to 474)

Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,

Lundberg, J., Uhlen, M. and Tunlid, A.

Transcriptional responses of *Paxillus involutus* and *Betula pendula*

during formation of ectomycorrhizal root tissue

Unpublished (2003)

Contact: Johansson, T.

Fungal-Hest Interaction Group (FHIG)

Microbial Ecology, Institution of Ecology

Ecology Building, Lund University, SE-223 62 Lund, Sweden

Tel: +46 46 222 45 49

Fax: +46 46 222 41 58

Email: tomas.johansson@mbic.kol.lu.se

PCR Primers

FORWARD: P104 (5'-GGGAGCGCGCATTTGTT-3')

BACKWARD: P105 (5'-AGTGAGCTCGAATTGGGCC-3')

Seq primer: P104

High quality sequence stop: 474.

Location/Qualifiers

1. 474

/organism="Paxillus involutus"

/mol_type="mRNA"

/strain="ATCC 200175"

/db_xref="taxon:71150"

/tissue_type="Mycelium"

/dev_stage="25 days of growth after transfer"

/lab_host="Escherichia coli BM25.8"

/clone_lib="PHIG:B) Axenic plate culture"

/note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This

EST clone is originating from one of three cDNA libraries,

constructed for transcript profiling of the mycorrhizal

interaction between the basidiomycete *Paxillus involutus*

and *Betula pendula* (birch). One library represents the

developed and functional mycorrhizal root tissue

(PHIG:A) Ectomycorrhiza plate culture', a second

library represents axenically grown fungus (PHIG:B)

Axenic plate culture') and a third library represents

axenically grown plants (PHIG:C) Axenic plate culture'.

Libraries were analyzed in parallel and 3555 (PHIG:A),

3964 (PHIG:B), and 2532 (PHIG:C) high-quality (PHRED 20)

ESTs of >99bp have been deposited. The cDNA libraries were

constructed from total RNA using the SMART cDNA library

construction kit (#K1051-1, Clontech, Palo Alto, CA, USA)

according to the manufacturer's instructions. Full-length

cDNAs were trimmed by SfiI, fractionated and directionally

ligated into (lambda)Triplex2 arms. The lambda library was

converted to a plasmid library via site-specific

recombination at loxP sites in a Cre+ strain (E. coli

BM25.8). Plasmid clones were randomly collected and

analysed by DNA sequencing using a plasmid-specific

forward primer (P104)."

ORIGIN

Query Match 4.1%; Score 79.2; DB 14; Length 474;

Best Local Similarity 54.6%; Pred. No. 0.0025;

Matches 203; Conservative 0; Mismatches 163; Indels 6; Gaps 2;

QY 1069 CAGCTTTGACATGGATGTTTCCACGACTCTCAGACGACAAAGCCGTTCCAGATAAATTC 1128

Db 79 CAGCGATTAACCTGGCGCTGATGATGTTGTCGTAGCCCTGGGATACAGAGAACTCC 138
QY 1129 GAGAGAAATTTGTCAGATCGACAGGATATGCTTACGCTAGACGAATTAATGCGTTGC 1188
Db 139 GGGAGAACTCTTCTCTGTGACACGGAACGCCCTGATGAGGAACTCTCGCGCTTC 198
QY 1189 CTATCTCGAAGCTTTGTTAGGAGTCTCTGCTAGACCCCTAGTCCGTTATGCTA 1248
Db 199 CTTACTTGACACCGTGTGAAGGAGACTTTGCGTGTACACCTTCTTTGGAGAGACCA 258
QY 1249 ACCGTGAATGCTTAAGAGTGAAGACTTTCATCCACTTTCGCGACCTGTCTATTTGGTCAG 1308
Db 259 TAGCGG---TCGCGATGAAGACGATATTTCTGCGCTCGAAGACCACTCAAGCAAAAC 315
QY 1309 ATGGTTCGCTATCAACAGGTCGCGATCAGAAAGNACGATGCTATGCTTCGTTGT 1368
Db 316 ATGGCGTGTACACGACGGAATTAGGATCAGAAAGGACCGCGATCTCTATCCCGATCT 375
QY 1369 TCAACATCAATCGTTCAAAAGTTTCATTTATGGAGAAGATGCAAGAAATTCAGACCGGAGA 1428
Db 376 T---GATGAACCGTTCGAAGGGTTGTGGGCCCTGACGCCACGAGTTCAAAACCGAAC 432
QY 1429 GTGGCTTGAGG 1440
Db 433 GCTGGACGACG 444

RESULT 3
CD275923
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD275923 475 bp mRNA linear EST 01-SEP-2003
T143B00226 (FHIG:B) Axenic plate culture Paxillus involutus cDNA
5', mRNA sequence.
CD275923
CD275923.1 GI:34387969
EST
Paxillus involutus
Paxillus involutus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Boletales; Paxillineae; Paxillaceae; Paxillus.
1 (bases 1 to 475)
Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
Lundberg, J., Uhlen, M. and Tunlid, A.
Transcriptional responses of Paxillus involutus and Betula pendula
during formation of ectomycorrhizal root tissue
Unpublished (2003)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@mbioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAGCGCGCCATTGTGT-3')
BACKWARD: P105 (5'-AGTAGCTCGAATTCGGCC-3')
Seq primer: P104
High quality sequence stop: 475.
Location/Qualifiers
1. .475
/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 200175"
/db_xref="taxon:71150"
/tissue_type="mycelium"
/dev_stage="25 days of growth after transfer"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(FHIG:B) Axenic plate culture"
/notes="Vector: pTriplex2; Site 1: Sfil; Site 2: Sfil; This
EST clone is originating from one of three cDNA libraries,
constructed for transcript profiling of the mycorrhizal
interaction between the basidiomycete Paxillus involutus
and Betula pendula (birch). One library represents the
developed and functional mycorrhizal root tissue

FEATURES
source

(('FHIG:A) Ectomycorrhiza plate culture'), a second library represents axenically grown fungus (('FHIG:B) Axenic plate culture') and a third library represents axenically grown plants (('FHIG:C) Axenic plate culture'). Libraries were analyzed in parallel and 3555 (FHIG:A), 3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20) ESTs of >99bp have been deposited. The cDNA libraries were constructed from total RNA using the SMART cDNA library construction kit (#K1051-1, Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. Full-length cDNAs were trimmed by Sfil, fractionated and directionally ligated into (lambda)tripleX2 arms. The lambda library was converted to a plasmid library via site-specific recombination at loxP sites in a Cre+ strain (E. coli RM25.8). Plasmid clones were randomly collected and analysed by DNA sequencing using a plasmid-specific forward primer (P104)."

Query Match 4.1%; Score 79.2; DB 14; Length 475;
Best Local Similarity 54.6%; Pred. No. 0.0025;
Matches 203; Conservative 0; Mismatches 163; Indels 6; Gaps 2;

QY 1059 CAGTCTTGACATGATGTTTACCGACTCTCAGAGACAAAGCCGTTTCAGGATAACTTC 1128
Db 88 CAGCGATAACTTGGCGCTGATGAGTTGTGCTAGCCCTGGGATACAGAGAGCTCC 147
QY 1129 GAGAAGAAATTTGTCAGATCGACACGGATATGCTTACGCTAGACGAATTAATTCGTTGC 1188
Db 148 GGGAGGAACCTTCTTCTGTTGACACGGAACGCCCTCGATGAGAGCACTCTCGCGCTTC 207
QY 1189 CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTGTCTAGACCTCTTCTAGTCCGTTGCTA 1248
Db 208 CTTACTTTCACACCGCTGTGTAAGGAGACTTTGCGTGTACACCTCTCCCTTTGGAGAGCA 267
QY 1249 ACCGTGAATGCTTAAAGATGAAGACTTTCATCCACTTGCAGGCTGTCTATGGTCGAG 1308
Db 258 TAGCGG---TCGCGATGAAGACGATATTTGCGCTCGAAGAGCCACTTCAGGACAAAC 324
QY 1309 ATGGTTCGCTCATCAACGAGTCCGATCATCAAGAAAGCAAGATGTTGCTCTCCGTTGT 1368
Db 325 ATGGCGTGTACACGCGGAATTAGGATCAGCAAGGACCGCGATCTCTATCCGATCT 384
QY 1369 TCAACATCAATCGTTCAAAAGTTTCATTTATGGAGAAGATGCAAGAAATTCAGACCGGAGA 1428
Db 385 T---GATGAACCGTTCGAAGGGTTGTGGGCCCTGACGCCACGAGTTCAAAACCGAAC 441
QY 1429 GTGGCTTGAGG 1440
Db 442 GCTGGACGACG 453

RESULT 4
AT002896
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

AT002896 615 bp mRNA linear EST 25-MAR-2002
AT002896 POSLM01 Pleurotus ostreatus cDNA clone 355LM, mRNA
sequence.
AT002896
AT002896.1 GI:6934623
EST.
Pleurotus ostreatus (oyster mushroom)
Pleurotus ostreatus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Pleurotaceae; Pleurotus.
1 (bases 1 to 615)
Lee, S.H., Kim, B.G., Kim, K.J., Lee, J.S., Yun, D.W., Hahn, J.H.,
Kim, G.H., Lee, K.H., Suh, D.S., Kwon, S.T., Lee, C.S. and Yoo, I.B.
Comparative Analysis of Sequences Expressed during the
Liquid-Cultured Mycelia and Fruit Body Stages of Pleurotus
ostreatus
Fungal Genet. Biol. 35 (2), 115-134 (2002)
2183865
11848675

COMMENT Contact: Kim BG
Division of applied microbiology
National Institute of Agricultural Science and Technology (NIAT)
Seodundeong, Suwon, Kyung-gi 441-707, South Korea
Email: bgkimniat.go.kr
Submitted through ERIC(Biological Research Information Center) of Korea
URL: http://bric.postech.ac.kr/.
Location/Qualifiers
1. .615
/organism="Pleurotus ostreatus"
/mol_type="mRNA"
/culivar="ASI 2029"
/db_xref="taxon:5322"
/clone="355IM"
/dev_stage="shaking liquid cultured mycelia"
/lab_host="E.coli"
/clone_lib="POSIM01"
/note="Vector: lambda Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; average insert size: 1000 bp; initial pfu: 5 X 10⁷
Library information: Isolation of total RNA from the mycelia incubated in shaking liquid MCM media at 30 deg C"

FEATURES
source
1. .615
/organism="Pleurotus ostreatus"
/mol_type="mRNA"
/db_xref="taxon:5322"
/clone="355IM"
/dev_stage="shaking liquid cultured mycelia"
/lab_host="E.coli"
/clone_lib="POSIM01"
/note="Vector: lambda Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; average insert size: 1000 bp; initial pfu: 5 X 10⁷
Library information: Isolation of total RNA from the mycelia incubated in shaking liquid MCM media at 30 deg C"

ORIGIN
Query Match 3.9%; Score 75.8; DB 9; Length 615;
Best Local Similarity 48.9%; Pred. No. 0.008;
Matches 234; Conservative 0; Mismatches 242; Indels 3; Gaps 1;
QY 1265 GGATGAAGACTTCATCCCACTTCGCGAGCCTGCTATGTCGAGATGGTGGTCAATCA 1324
DB 59 GGTGATGATGTTCTCCCGTGGAGCGTCCGTTGACAGATCTGAAGGGTATTACTCATCA 118
QY 1325 CGAGGTCCGGATCAGGAAGAACGATGCTATGCTTCGCTGTTGTTCAACATCAATCGTTC 1384
DB 119 ATCAGTAGCAATCAAGAAGAACATGTTTATGTTCCGATCTTCGGATGATGATTT 178
QY 1385 AAAGTTCAATTTATGAGAGATGCGAAGATTCAGACCGGAGAGTGGCTTGAGGACGT 1444
DB 179 GGTATCAATATGGGCGAGGATGCTTTTGAATTCAGCCAGACAGCTTGGC---AATCACC 235
QY 1445 AACAGACTCGCTACACGATTTGAGACCCCTATGGACACGAGCGAGCTTTATCTCTGG 1504
DB 236 GCGGAGGCTGCAGCGTGTCCAGGATATGAGATGAGTAACTGATGATGTTCTCTGTTGG 295
QY 1505 ACCAGAGCTGCTTTGTTGGCGATTTGCTGTCCGAGATGAAGCGCTTTGTTGTTG 1564
DB 296 TCCTCGAGCATGATGCTGTTGAGATTTCTATCGTAGATGAAGCGCTTGTGTTTCCAC 355
QY 1565 CACTCTCCGTCGGTCCAGTTCGAGCCCATCTCTCTATCCAGATACGAGCACATCAC 1624
DB 356 CTTGTGCGTTCATTGAGTTCGAGTTCGCTGTTCCATCAGAGAAATTTGGCAAGATCTC 415
QY 1625 CTTGATCATTTCCGCTCCGATCGTGGTAGAGAGAGAGGGGTACAGATGCGTTT 1684
DB 416 GTCCATCGTCGAGGACCTATTCTGAAGAACGATAAGAGCGGGAATGTGATGCCCTT 475
QY 1685 GCAGGTCAAGCGGTGGAAGTGTGATCTTCTATATGTTTAAGAGAAGTCTATATCTG 1743
DB 476 GATCATTAACCATATATCAACGGAGGTGACCTCAACATTTAAACGATGATATGATGCG 534

RESULT 5
CD274181
LOCUS
DEFINITION
Ti43B02B39F (FHIG:B) Axenic plate culture Paxillus involutus cDNA
5', mRNA sequence.
ACCESSION
CD274181
VERSION
CD274181.1 GI:34386227
KEYWORDS
EST.
SOURCE
Paxillus involutus
Paxillus involutus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Boletales; Paxillaceae; Paxillus.

REFERENCE 1 (bases 1 to 354)
Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
Lundeberg, J., Uhlen, M. and Tunlid, A.
Transcriptional responses of Paxillus involutus and Betula pendula
during formation of ectomycorrhizal root tissue
Unpublished (2003)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@bioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAAGCGCGCATTTGGTT-3')
BACKWARD: P105 (5'-AGTGAGCTGGAATGGCGCC-3')
Seq primer: P104
High quality sequence stop: 354.
Location/Qualifiers
1. .354
/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 200175"
/db_xref="taxon:71150"
/tissue_type="Mycelium"
/dev_stage="25 days of growth after transfer"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(FHIG:B) Axenic plate culture"
/note="Vector: pTriplex2; Site 1: SfiI; Site 2: SfiI; This
EST clone is originating from one of three cDNA libraries,
constructed for transcript profiling of the mycorrhizal
interaction between the basidiomycete Paxillus involutus
and Betula pendula (birch). One library represents the
developed and functional mycorrhizal root tissue
(('FHIG:A) Ectomycorrhiza plate culture'), a second
library represents axenically grown fungus ('(FHIG:B)
Axenic plate culture') and a third library represents
axenically grown plants ('(FHIG:C) Axenic plate culture').
Libraries were analyzed in parallel and 3555 (FHIG:A),
3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)
ESTs of >99bp have been deposited. The cDNA libraries were
constructed from total RNA using the SMART cDNA library
construction kit (#K1051-1, Clontech, Palo Alto, CA, USA).
According to the manufacturer's instructions, full-length
cDNAs were trimmed by SfiI, fractionated and directionally
ligated into (lambda)Triplex2 arms. The lambda library was
recombined at loxP sites in a Cre+ strain (E. coli
BM25.8). Plasmid clones were randomly collected and
analysed by DNA sequencing using a plasmid-specific
forward primer (P104)."

ORIGIN
Query Match 3.6%; Score 69.2; DB 14; Length 354;
Best Local Similarity 54.4%; Pred. No. 0.12;
Matches 162; Conservative 0; Mismatches 133; Indels 3; Gaps 1;
QY 1069 CAGCTTGACATGATGTTTACCGACTCTCAGAAGACAAAGCGTTCAGGATAAATCTC 1128
DB 60 CAGCGATAACTTGGCGCGCTGCATGAGTTGTGCTAGCCCTGGGATACAGAGAAGCTCC 119
QY 1129 GAGAAGAAATTTGTTCAGATCGACATGATGCTTACGCTAGACGAACATTTAATCGCTTC 1188
DB 120 GGGAGGAATCTTCTTCTGTTGACACGGAAGCGCCCTCGATGGACGAATCTCGCGGCTTC 179
QY 1189 CTTATCTCGAAGCGTTTGTTHAAGAGTCTCTTGTCTAGACCCCTCTAGTCCGATGCTA 1248
DB 180 CCTACCTTGACACCGTGTGTAAGAGAGATTTGGCGTGTACACCCCTCCCTTTGTAGAGACCA 239
QY 1249 ACCGTGAATGTTAAAGGATGAAGACTTCATCCACCTTGGCGAGCTGTCAATGGTCGAG 1308
DB 240 TAGCGG---TCGCGATGAAGACCATATTTCTGCGCTCGAGAGGCCACTCAGCAAAAC 296

QY	1069	CAGTCTTGACATGATGTTTCACCGACTCTCAGAGAGACAAAGCCGCTTCAGGATAAACTTC	1122
DB	99	CAGCGATTAACCTTGGCCCTCGATGAGTTGTGCTAGCCCTGGGATACAGAGGAGCTCC	158
QY	1129	GAGAGAGAAATTTGTCCAGATCGACACGGATATGCCTACGGTAGACGAACTTAATCGGTTGC	1188
DB	159	GGGAGGAACCTCTTCTTGTTGACACGGAAACGCCCTCGATGGACGAACTCTCGCGCTTC	218
QY	1189	CTTATCTCGAAGCGTTGTTAAGGAGTCTCTTGCTAGACCCCTCCTAGTCGGTAGTGTA	1248
DB	219	CCTACCTTGAACACCGTGGGAGAGAGACTTTGGTGATACACCTCCCTTTGGAGNAGCA	278
QY	1249	ACCGTGAATGCTTAAAGATGAAGACTTCATCCCACTTCGCCAGCGCTGCTATTTGGTCGAG	1308
DB	279	TACGCG---TCGCGATGAAGACGATATTCTGCGCTCGAGAAGCCACTCACGGACAAC	335
QY	1309	ATGGGTCCGTCAATCAACGAGGTCGGATCACGAAGGAACGATGCTCATGCTCCGTTGT	1368
DB	336	ATGGCGGTGCTACACGACGGAATTAGGATCAGCAGAGGACCGGATCTCTATCCGATCT	395
QY	1369	TCAACATC	1376
DB	396	TGATGAAC	403
RESULT 7			
BQ102580			
LOCUS			
DEFINITION			
BQ102580			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. .487			
/organism="Schizopyllum commune"			
/mol_type="mRNA"			
/strain="4-40"			
/db_xref="taxon:5334"			
/tissue_type="mycelium"			
/clone_lib="MIN Nitrogen-replete Schizopyllum library"			
/note="Vector: lambda Zap; Site 1: EcoRI; Site 2: XhoI; 4-day-old mycelia of Schizopyllum commune were transferred from minimal (nitrogen-replete) medium to fresh minimal medium. RNA was extracted twelve hours after transfer and cDNAs prepared."			
POLYA=NO.			
Location/Qualifiers			
1. .487			
/organism="Schizopyllum commune"			
/mol_type="mRNA"			
/strain="4-40"			
/db_xref="taxon:5334"			
/tissue_type="mycelium"			
/clone_lib="MIN Nitrogen-replete Schizopyllum library"			
/note="Vector: lambda Zap; Site 1: EcoRI; Site 2: XhoI; 4-day-old mycelia of Schizopyllum commune were transferred from minimal (nitrogen-replete) medium to fresh minimal medium. RNA was extracted twelve hours after transfer and cDNAs prepared."			
POLYA=NO.			
Location/Qualifiers			
1. .487			
/organism="Schizopyllum commune"			
/mol_type="mRNA"			
/strain="4-40"			
/db_xref="taxon:5334"			
/tissue_type="mycelium"			
/clone_lib="MIN Nitrogen-replete Schizopyllum library"			
/note="Vector: lambda Zap; Site 1: EcoRI; Site 2: XhoI; 4-day-old mycelia of Schizopyllum commune were transferred from minimal (nitrogen-replete) medium to fresh minimal medium. RNA was extracted twelve hours after transfer and cDNAs prepared."			
POLYA=NO.			
Location/Qualifiers			
1. .487			
/organism="Schizopyllum commune"			
/mol_type="mRNA"			
/strain="4-40"			
/db_xref="taxon:5334"			
/tissue_type="mycelium"			
/clone_lib="MIN Nitrogen-replete Schizopyllum library"			
/note="Vector: lambda Zap; Site 1: EcoRI; Site 2: XhoI; 4-day-old mycelia of Schizopyllum commune were transferred from minimal (nitrogen-replete) medium to fresh minimal medium. RNA was extracted twelve hours after transfer and cDNAs prepared."			
POLYA=NO.			
Location/Qualifiers			
1. .487			
/organism="Schizopyllum commune"			
/mol_type="mRNA"			
/strain="4-40"			
/db_xref="taxon:5334"			


```
Db 1 GCATTGGATGACCGGTGCTACCTTTGATGCTCTCGGTGTCGACGGCTTCGACTACAACTT 60
QY 674 CGACTCGCTCCAGACACAGACCAATGAGTCTATGTCGCTTTTGTGCGACTTACCGATGG 733
Db 61 CGAGGCGATCCACACGAGGACATGAGTCTCAATGCGTATTAAGACATGTCGAACT 120
QY 734 GTTTGCTCTCACTTGGAGTCGTTCAAGGCTATCATGTTGGGATTTTGTACCTTACTTCGG 793
Db 121 GGCTGTGTGCGAAGCGAGCTTCTGCGACAGCATGCAACAACTTTACGCGCGCTTCTGCT 180
QY 794 AACTATGAACGGAGACATGAGATAC---CTTGAAGTCAAGGATTAAGCAAGTTTCCCGACG 850
Db 181 AGCCTCTTCCCTACCCAGAACAGCGGACCGTGAACGCTGCCGAGAGTCAATCCGGCG 240
QY 951 AGTGGGATCGAGCTTATGAGCAAAAGACAGCGCGTGTGTTGGCTCAGCTTCCGATCA 910
Db 241 CGTACGCGGCGACATCATCCAGCACACAGAAGC-----GCAAGATCGAGGAGGG 288
QY 911 GGCTGTGATAAAAGGATGTTCAAGGTCGGGATATCTTAAGTCTCCTAGTGAGAGCAA 970
Db 289 CATGCGCGAGCGCAACCTTAGCGGCCAAGACATCTCACCTCTCTCTCAAGTCCAA 348
QY 971 CATGCGCGCAACCTGCTCCCTGAATCTCAAAAGCTGTCCGATGAGGAGTACTCGCTCAG 1030
Db 349 CATGCTCCACCGACATCCCGCCCGACCGCGCATCACCGACGCCACCTCTCTCGACAAC 408
QY 1031 CAGTAACCTGTTATTTGCTGGATATGAACCTTCTCGACAGTCTTGATGAGATTTCA 1090
Db 409 CAACAGCTGCGCTTGGCGGGTCCGACACCTCGTCCCTCGGCTCACTCGGACGCTCCA 468
QY 1091 CCGACTCTCAGAAGAC 1106
Db 469 CTTCTCGCGAGCAC 484
```

```
RESULT 8
CD275427
LOCUS
DEFINITION T143B04144F (FHIG:B) Axenic plate culture Paxillus involutus cDNA
ACCESSION CD275427
VERSION 1
KEYWORDS 5', mRNA sequence.
SOURCE CD275427.1 GI:34387473
ORGANISM Paxillus involutus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Boletales; Paxillineae; Paxillaceae; Paxillus.
REFERENCE 1 (bases 1 to 400)
AUTHORS Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
Lundeberg, J., Uhlen, M. and Tunlid, A.
TITLE Transcriptional responses of Paxillus involutus and Betula pendula
JOURNAL during formation of ectomycorrhizal root tissue
COMMENT Unpublished (2003)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@bioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAAGCGCGCATTTGTTT-3')
BACKWARD: P105 (5'-AGTGAGCTCGAATTCGGCC-3')
Seq primer: P104
High quality sequence stop: 400.
Location/Qualifiers
1. 400
/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 500175"
/db_xref="taxon:71150"
/tissue_type="Mycelium"
```

FEATURES

source

```
/dev_stage="25 days of growth after transfer"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(FHIG:B) Axenic plate culture"
/notes=vector: Triplex2; Site: 1: Sfil; Site: 2: Sfil; This
EST clone is originating from one of three cDNA libraries,
constructed for transcript profiling of the mycorrhizal
interaction between the basidiomycete Paxillus involutus
and Betula pendula (birch). One library represents the
developed and functional mycorrhizal root tissue
('FHIG:A) Ectomycorrhiza plate culture'), a second
library represents axenically grown fungus ('FHIG:B)
Axenic plate culture') and a third library represents
axenically grown plants ('FHIG:C) Axenic plate culture').
Libraries were analyzed in parallel and 3555 (FHIG:A),
3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20)
ESTs of >99bp have been deposited. The cDNA libraries were
constructed from total RNA using the SMART cDNA library
Construction kit (#1051-1, Clontech, Palo Alto, CA, USA)
according to the manufacturer's instructions. Full-length
cDNAs were trimmed by Sfil, fractionated and directionally
ligated into (lambda)Triplex2 arms. The lambda library was
recombination at loxp sites in a Cre+ strain (E. coli
BM25.8). Plasmid clones were randomly collected and
analysed by DNA sequencing using a plasmid-specific
forward primer (P104)."
```

ORIGIN

```
Query Match 3.0%; Score 58.8; DB 14; Length 400;
Best Local Similarity 59.6%; Pred. No. 5.9; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 67;

QY 1069 CAGTCTTGACATGATGTTTCACCGACTCTCAGAAGCAAGCGCTTCAGATPAACTTC 1128
Db 52 CAGCGATTAACCTGGCGCTGCATGATGTCGTAGCCCTCGGATACAGAGGAGCTCC 111
QY 1129 GAGAAGAAATTTGTCAGATCGACGGATATGCTAGCTAGACGACTTAATGCGTTGC 1188
Db 112 GCGAGGAACTCTCTTCTGTTGACAGGAAAGCGCCCTCGATGGACGAACTCTCGCGGCTTC 171
QY 1189 CTTATCTCGAAGCGTTGTTTAAGGAGTCTCTGCTAGACCCCTCC 1234
Db 172 CTTACTTGCACCGTGGTGAAGGAGACTTGGGTGTACACCCCTCC 217
```

RESULT 9

```
CD275197
LOCUS
DEFINITION CD275197 317 bp mRNA linear EST 01-SEP-2003
5', mRNA sequence.
ACCESSION CD275197
VERSION 1
KEYWORDS GI:34387243
SOURCE CD275197.1
ORGANISM Paxillus involutus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Boletales; Paxillineae; Paxillaceae; Paxillus.
REFERENCE 1 (bases 1 to 317)
AUTHORS Johansson, T., Le Quere, A., Ahren, D., Soderstrom, B., Erlandsson, R.,
Lundeberg, J., Uhlen, M. and Tunlid, A.
TITLE Transcriptional responses of Paxillus involutus and Betula pendula
JOURNAL during formation of ectomycorrhizal root tissue
COMMENT Unpublished (2003)
Contact: Johansson, T.
Fungal-Host Interaction Group (FHIG)
Microbial Ecology, Institution of Ecology
Ecology Building, Lund University, SE-223 62 Lund, Sweden
Tel: +46 46 222 45 49
Fax: +46 46 222 41 58
Email: tomas.johansson@bioekol.lu.se
PCR Primers
FORWARD: P104 (5'-GGGAAGCGCGCATTTGTTT-3')
BACKWARD: P105 (5'-AGTGAGCTCGAATTCGGCC-3')
Seq primer: P104
High quality sequence stop: 400.
Location/Qualifiers
1. 400
/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 500175"
/db_xref="taxon:71150"
/tissue_type="Mycelium"
```

Seq primer: P104
High quality sequence stop: 317.
Location/Qualifiers
1. 317

FEATURES
source

/organism="Paxillus involutus"
/mol_type="mRNA"
/strain="ATCC 200175"
/db_xref="taxon:71150"
/tissue_type="mycelium"
/dev_stage="25 days of growth after transfer"
/lab_host="Escherichia coli BM25.8"
/clone_lib="(FHIG:B) Axenic plate culture"
/note="(FHIG:B) Site 1: Sfil; Site 2: Sfil; This EST clone is originating from one of three cDNA libraries, constructed for transcript profiling of the mycorrhizal interaction between the basidiomycete Paxillus involutus and Betula pendula (birch). One library represents the developed and functional mycorrhizal root tissue ('(FHIG:A) Ectomycorrhiza plate culture'), a second library represents axenically grown fungus ('(FHIG:B) Axenic plate culture'), and a third library represents axenically grown plants ('(FHIG:C) Axenic plate culture'). Libraries were analyzed in parallel and 3555 (FHIG:A), 3964 (FHIG:B), and 2532 (FHIG:C) high-quality (PHRED 20) ESTs of >99bp have been deposited. The cDNA libraries were constructed from total RNA using the SMART cDNA library construction kit (#K1051-1, Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. Full-length cDNAs were trimmed by Sfil, fractionated and directionally ligated into (lambda)Triplex2 arms. The lambda library was converted to a plasmid library via site-specific recombination at loxp sites in a Cre+ strain (E. coli BM25.8). Plasmid clones were randomly collected and analysed by DNA sequencing using a plasmid-specific forward primer (P104)."

ORIGIN

Query Match 3.0%; Score 58; DB 14; Length 317;
Best Local Similarity 55.4%; Pred. No. 8.7;
Matches 112; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 1069 CAGTCTTGACATGGATGTTTCACCGACTCTCAGAGACAAAGCCGCTTCAGGATAAATTC 1128
Db 60 CACGGATACTTGGCGCTGCTAGTGTCTGCTAGCCCTGGGATACAGAGAGTCC 119

QY 1129 GAGAGAAATTTGTAGATCGACAGATATGCCATGCTAGACGAACTTAATCGTTGC 1188
Db 120 GGAGGAACTTCTTCTGTTGACACGAAACGCCCTCGATGGACGAACCTCGCGCTTC 179

QY 1189 CTTATCTCGNAGCGTTTGTAGGAGTCTCTTCGCTAGACCTCCTAGTCCGTATGCTA 1248
Db 180 CTTACCTTGACCGGTGGTGAAGGACATTTGCGTGTGCACCTCCCTTTGGAGAGACCA 239

QY 1249 ACCGTGAATCTTTAAAGGATGA 1270
Db 240 TACGCGTCCGATGAAGACGA 261

RESULT 10
BJ075377
LOCUS
DEFINITION BJ075377 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XL057f08 5', mRNA sequence.
ACCESSION BJ075377.1 GI:17520293
VERSION BJ075377.1
KEYWORDS EST.
ORGANISM Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 499).
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and

Kohara,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following URL.
http://xenopus.nibb.ac.jp.

FEATURES
source

Location/Qualifiers
1. 499
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone_lib="XL057f08"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN

Query Match 2.9%; Score 56.8; DB 12; Length 499;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 143; Conservative 0; Mismatches 117; Indels 6; Gaps 1;

QY 996 CAAAAGCTGTCGGATGAGGAGTACTCGCTCAGATCACTGTAACCTGTTATTGCTGATAT 1055
Db 52 CAGAGCTTTCTGATGAGGATATACGGCTGAGTGACACATTCATGTTTGAGGACAT 111

QY 1056 GAAATTTCTTCGACAGTCTTCAGATGGATGTTTCCAGACTCTCAGAGACAAAGCGTT 1115
Db 112 GATACATCATCTAGTGAATATCTCTGGATCTCATATTGTATGGTACATCCAGAGCAC 171

QY 1116 CAGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCCTAGCTAGAC--- 1172
Db 172 CAGCAGAAATGCCAGGAGAGATCAATGAGTCTTGGGGAGCGACAAACCATGGATGG 231

QY 1173 ---GAACCTTAATGCGTTCCCTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGAC 1229
Db 232 GATGACCTTAATAGGATACCTTATACCACATGTGTATTAAAGAAAGTCTGCGCTATAC 291

QY 1230 CCTCCTAGTCCGTATGCTAACCGTGA 1255
Db 292 CCTCCAGTCCCATCATGATCCAGGA 317

RESULT 11
BX779027/c

LOCUS BX779027 XEC-egg Silurana tropicalis cDNA clone TEGG070n12 3', mRNA sequence.
DEFINITION BX779027.1 GI:39686233
ACCESSION BX779027
VERSION BX779027
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Silurana.

REFERENCE 1 (bases 1 to 971)
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
CONTACT: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TEGG070n12.q1kt7

Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dT primed from 5' of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1. .971
/organism="Silurana tropicalis"
/mol_type="rRNA"
/db_xref="taxon:8364"
/clone="TEG9070n12"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5' of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

FEATURES
source

ORIGIN

Query Match 2.9%; Score 55.4; DB 13; Length 971;
Best Local Similarity 54.7%; Pred. No. 15;
Matches 135; Conservative 0; Mismatches 106; Indels 6; Gaps 1;
QY 995 TCAAAGCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACTGTTATTGCTGGATA 1054
DB 817 TAAAGCTCTGACAGATGCAGAAATAATGCTCAGTCCCTTATATTTATATTGCTGGATA 758
QY 1055 TGAACCTCTTCGACAGTCTTCCATGATGTTTCCACGACTCTCAGAAACAAAGCCGT 1114
DB 757 TGAACCAACAGTACAGCACTCTCTTACCTGTTTATATCTTGCACCCACCTGACGT 698
QY 1115 TCAGGATAAAGCTCGAAGAAATTT-----TGTGATCGACAGGATGCTTACGCT 1168
DB 697 CCAGCAGAACTGCATGAGGATAGATTCATTTCTCCCTGACAGGCCAGTCTACTTA 638
QY 1169 AGACGAACCTTAATGGTTCCTTATCTCGAAGCGTTTGTAAAGAGTCTTCTGCTAGA 1228
DB 637 TGACATTTAATGCAGATGGAATATCTTGATATGTTTATTCAGGAGCCCTCAGTTGTT 578
QY 1229 CCTCCT 1235
DB 577 TCCCTCT 571

RESULT 12

CF152584
LOCUS
DEFINITION
AGENCOURT.14907205 NICHD XGC Emb8 Silurana tropicalis cDNA clone
IMAGE:6985215 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 849)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs@mail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LIA4652 row: h column: 14
High quality sequence start: 66
High quality sequence stop: 708.
Location/Qualifiers
1. .849
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6985215"
/tissue_type="tadpole"
/dev_stage="embryo, stages 40-45"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb8"
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.1 kb. Constructed by invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

FEATURES
source

ORIGIN

Query Match 2.8%; Score 53.8; DB 14; Length 849;
Best Local Similarity 54.3%; Pred. No. 29;
Matches 134; Conservative 0; Mismatches 107; Indels 6; Gaps 1;
QY 995 TCAAAGCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACTGTTATTGCTGGATA 1054
DB 480 TAAAGCTCTGACAGATGCAGAAATAATGCTCAGTCCCTTATATTTATATTGCTGGATA 539
QY 1055 TGAACCTCTTCGACAGTCTTGCATGATGTTTCCACGACTCTCAGAAACAAAGCCGT 1114
DB 540 TGAACCAACAGTACAGCACTGTTTACCTGTTTATATCTTGCACCCACCTGACGT 599
QY 1115 TCAGGATAAAGCTCGAAGAAATTT-----TGTGATCGACAGGATGCTTACGCT 1168
DB 600 CCAGCAGAGATGCTGATGAGGATAGATTCATTTCTCCCTGACAGGCCAGTCTACTTA 659
QY 1169 AGACGAACCTTAATGGTTCCTTATCTCGAAGCGTTTGTAAAGAGTCTTCTGCTAGA 1228
DB 660 TGACATTTAATGCAGATGGAATATCTTGATATGTTTATTCAGGAGCCCTTAGTTGTT 719
QY 1229 CCTCCT 1235
DB 720 TCCCTCT 726

RESULT 13

AW600900
LOCUS
DEFINITION
7C2 cDNA library of 4-day-old Eucalyptus globulus
bicostata-Pisolithus tinctorius ectomycorrhiza Pisolithus
tinctorius cDNA 5' similar to hydrophobin, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pisolithus tinctorius
Pisolithus tinctorius
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Boletales; Sclerodermatineae; Pisolithaceae; Pisolithus.
1 (bases 1 to 568)
Voiblet, C., Duplessis, S., Encelot, N. and Martin, F.
Identification of symbiosis-regulated genes in Eucalyptus
globulus-Pisolithus tinctorius ectomycorrhiza by differential
hybridization of arrayed cDNAs
Plant J. 25 (2), 181-191 (2001)
21097016
MEDLINE
PUBMED
COMMENT
Contact: Martin FM
Equipe de Microbiologie Forestiere
Institut National de la Recherche Agronomique
Centre INRA de Nancy, 54280 Champenoux, France

Tel: +33 383 39 40 80
 Fax: +33 383 39 40 69
 Email: fmartin@nancy.inra.fr
 Insert Length: 568 Std Error: 0.00
 Seq primer: Capfinder 5'CDS
 POLYA=Yes.

Location/Qualifiers

1..568
 /organism="Pisolithus tinctorius"
 /mol_type="mRNA"
 /strain="441"
 /db_xref="taxon:37468"
 /tissue_type="mycelium, 4 days after contact"
 /dev_stage="mycelium, 4 days after contact"
 /lab_host="Eucalyptus globulus bicostata"
 /clone_lib="cDNA library of 4-day-old Eucalyptus globulus bicostata-Pisolithus tinctorius ectomycorrhiza"
 /note="Organ: ectomycorrhiza; Vector: pBluescript; Site_1: EcoRI; The cDNA library was constructed from 1 ug total ectomycorrhiza RNA using the SMART cDNA library construction kit (#K1051-1, Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. cDNAs were cloned into EcoRI-digested pBluescript."

FEATURES
source

Query Match 2.7%; Score 51.2; DB 10; Length 568;
 Best Local Similarity 58.6%; Pred. No. 91;
 Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1762 AATGCTCTTTGATCGATTTGTTCTCATACCGGCGCGCTATGACTTCTAGTC 1821
 |||||
 Db 400 AATACATTTTCATTTTGGTGTGTCGGGACAGATGGAATTTGATTTGTTCTACTACGTA 459
 |||||

QY 1822 GTCATCGTCGCTCGACTCTCTTCTACCTATATATTCATCCGAAAGAAAAA 1881
 |||||
 Db 460 TTGCACGGTATAGGAGTCTAGCTTCTATTATAAATGCTGTTCTCTAAAAAAA 519
 |||||

QY 1882 AAAAAAAGGCGCGCTCGA 1913
 |||||
 Db 520 AAAAAAAGGCGCGCTCGA 551
 |||||

RESULT 14

AW600865 1296 bp mRNA linear EST 23-MAR-2000
 LOCUS
 DEFINITION
 5C8 cDNA library of 4-day-old Eucalyptus globulus
 bicostata-Pisolithus tinctorius ectomycorrhiza Pisolithus
 tinctorius cDNA 5' similar to cross-pathway control wd-repeat
 protein cpc-2, mRNA sequence.

ACCESSION AW600865.1 GI:7304838
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Pisolithus tinctorius
 Pisolithus tinctorius
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 Boletales; Sclerodermatineae; Pisolithaceae; Pisolithus.

REFERENCE
 AUTHORS
 TITLE
 1. (bases 1 to 1296)
 Voiblet C., Duplessis S., Encelet N. and Martin F.
 Identification of symbiosis-regulated genes in Eucalyptus
 globulus-Pisolithus tinctorius ectomycorrhiza by differential
 hybridization of arrayed cDNAs

JOURNAL
 MEDLINE
 PUBMED
 Plant J. 25 (2), 181-191 (2001)
 11169194

COMMENT
 Contact: Martin FM
 Equipe de Microbiologie Forestiere
 Institut National de la Recherche Agronomique
 Centre INRA de Nancy, 54280 Champenoux, France
 Tel: +33 383 39 40 80
 Fax: +33 383 39 40 69
 Email: fmartin@nancy.inra.fr
 Insert Length: 1296 Std Error: 0.00
 Seq primer: Capfinder 5'CDS

POLYA=Yes.

FEATURES
 source

Location/Qualifiers

1..1296
 /organism="Pisolithus tinctorius"
 /mol_type="mRNA"
 /strain="441"
 /db_xref="taxon:37468"
 /tissue_type="mantle and Hartig net hyphae"
 /dev_stage="symbiotic, 4 days after contact"
 /lab_host="Eucalyptus globulus bicostata"
 /clone_lib="cDNA library of 4-day-old Eucalyptus globulus bicostata-Pisolithus tinctorius ectomycorrhiza"
 /note="Organ: ectomycorrhiza; Vector: pBluescript; Site_1: EcoRI; The cDNA library was constructed from 1 ug total ectomycorrhiza RNA using the SMART cDNA library construction kit (#K1051-1, Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. cDNAs were cloned into EcoRI-digested pBluescript."

ORIGIN

Query Match 2.6%; Score 50.6; DB 10; Length 1296;
 Best Local Similarity 76.5%; Pred. No. 82;
 Matches 62; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1833 CTCGTGACTCTCTTACCTATATATTCATCCGAAAGAAAAA 1892
 |||||
 Db 1150 CTCGCCATATTCGTTGGGTCATAGCCATCCCGGAAAAA 1209
 |||||

QY 1893 AAAAAAAGCGCGCTCGA 1913
 |||||
 Db 1210 AAAAAAAGCGCGCTCGA 1230
 |||||

RESULT 15

CB200527 875 bp mRNA linear EST 21-MAR-2003
 LOCUS
 DEFINITION
 AGENCOURT 11243932 NICHD_XGC_Tad2 Xenopus laevis cDNA clone
 IMAGE:6875638 5', mRNA sequence.

ACCESSION CB200527
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Xenopus laevis (African clawed frog).
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.

REFERENCE
 AUTHORS
 TITLE
 1. (bases 1 to 875)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

JOURNAL
 COMMENT
 On Feb 4, 2003 this sequence version replaced gi:28233133.
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Drs. Donald Brown and Liqian Cai
 cDNA Library Preparation: CLONTECH
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LNC3111 row: f column: 21
 High quality sequence stop: 693.
 Location/Qualifiers
 1..875

FEATURES
source

/organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:6875638"
 /dev_stage="metamorphosis stage 62"
 /clone_lib="NICHD_XGC_Tad2"
 /note="Organ: Developing Tadpole; Vector: pDNR-LIB;
 Site_1: Sfil; Site_2: Sfil; 5' and 3' adaptors were used

in cloning as follows: 5' adaptor sequence:
5'-CACGCGCATATGGCC-3' and 3' adaptor sequence:
5'-ATCTAGAGCGCGGCGGCGGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.7 kb
(range 0.8-3.0 kb). 15/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match	2.6%	Score 50.4;	DB 14;	Length 875;
Best Local Similarity	52.3%	Pred. No. 1e+02;		
Matches 139;	Conservative 0;	Mismatches 121;	Indels 6;	Gaps 1;

QY	996	CAAAAGCTGTCGATGAGGAGGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATAT	1055
DB	101	CAAGGGCTTCTGATGAGGATATACGGCTGAGTGGACACATTTCATGTTTGGGGACAT	160
QY	1056	GAACTTCTTCGACAGTCTTGACATGGATGTTTCCCGACTCTCAGAGACAAGCCGTT	1115
DB	161	GATACTACATCTAGTGAATATCCTGGATCCTATATTGTTATGGCTACACATCCAGAGCAC	220
QY	1116	CAGGATAAATTCGAGAGAAATTTGTGAGATCGACACGGATATGCTAGCTAGAC---	1172
DB	221	CAGCAGAAATGCCAGAGGAGATCAACGAAGTCTTGGGGAGCGACAACCATGGATGG	280
QY	1173	---GAACTTAATGCGTTGCGCTTATCTCGAAGCGTTTGTTAAGGAGTCTCTTCTCTAGAC	1229
DB	281	GATGACCTTGGTAGGATACCTTATACCAATGTGTATTAGGAAGTCTGCGCCTATAC	340
QY	1230	CCTCCTAGTCGGTATGCTAACCGTGA	1255
DB	341	CCTCCAGTCCCATCAGTATCCAGGA	366

Search completed: April 6, 2004, 17:46:05
Job time : 5202 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 08:11:57 ; Search time 7663 Seconds
(without alignments)
10927.671 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932

Sequence: 1 gaattcgacagagccacc.....agccggctcgtccgaattc 1932

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pin.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1932	100.0	1932	6	AR203449 Sequence
2	1932	100.0	1932	6	AX034665 Sequence
3	353.2	18.3	3969	6	AR203450 Sequence
4	353.2	18.3	3969	6	AX034666 Sequence
5	55	2.8	2514	5	BC060496 Xenopus l
6	51.8	2.7	2000	6	AX655393 Sequence
7	48.4	2.5	1062	9	HSAS53377 Homo sapi
8	48.4	2.5	1365	9	HSAS53376 Homo sapi
9	48.4	2.5	1512	6	E10636 Human cDNA
10	48.4	2.5	1512	6	E10855 cDNA encodi
11	48.4	2.5	1512	6	AR399323 Sequence
12	48.4	2.5	1512	9	HSAS53375 Homo sapi
13	48.4	2.5	2011	6	AX408939 Sequence
14	48.4	2.5	2011	9	HUMCYPHLP
15	48.4	2.5	2011	9	HUMPA450M
16	48.4	2.5	2036	9	HUMCIPNO
17	48.4	2.5	2036	9	AR380517 Sequence
18	48.4	2.5	2059	9	HUMCYPNOA
19	48.4	2.5	2059	11	G18329
20	48.4	2.5	2759	6	AR222892 Sequence
21	48.4	2.5	2759	9	HSRCYP3
22	48.4	2.5	2849	6	AX409463 Sequence
23	48.4	2.5	2849	9	HUMCIT111A
24	48	2.5	15452	2	AC137912
25	47.4	2.5	1410	8	AB084385
26	47.4	2.5	1684	10	BC040779
27	47.4	2.5	1871	8	AK072295
28	47.4	2.5	1900	8	AK101670
29	47.4	2.5	2160	6	AR403280
30	47.4	2.5	2160	6	BD140625
31	47.2	2.4	3588	9	HSM802481
32	46.8	2.4	1512	9	AY334551
33	46.8	2.4	1704	9	AF182273
34	46.8	2.4	1712	6	AR352699
35	46.8	2.4	1712	6	BD195644
36	46.8	2.4	1822	6	AR352698
37	46.8	2.4	1822	6	BD195643
38	46.8	2.4	2032	9	SS3047
39	46.8	2.4	2139	9	BC015814
40	46.8	2.4	6004	10	AF372834
41	46.6	2.4	1395	5	AB091340
42	46.4	2.4	1760	3	AY051533
43	46.2	2.4	1286	9	AF131809
44	46.2	2.4	179437	5	AL928892
45	46.2	2.4	197330	2	AL928858

ALIGNMENTS

RESULT 1	AR203449	1932 bp	DNA	linear	PAT 20-JUN-2002
LOCUS	AR203449	Sequence 2 from patent US 6365386.			
DEFINITION	Sequence 2 from patent US 6365386.				
ACCESSION	AR203449				
VERSION	AR203449.1	GI:21499840			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1932)				
AUTHORS	Hoshino, T., Ojima, K. and Setoguchi, Y.				
TITLE	Astaxanthin synthase				
JOURNAL	Patent: US 6365386-A 2 02-APR-2002;				
FEATURES	Location/Qualifiers				

```

source      1. .1932
            /organism="unknown"
            /mol_type="unass-gned DNA"

ORIGIN
Query Match      100.0%; Score 1932; DB 6; Length 1932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAATTCGGCAGCGCCACCTACTCTTCTCCATATGTTTCAATCTCTTGGCTTGTCTCAAGGTG 60
Db 1 GAATTCGGCAGCGCCACCTACTCTTCTCCATATGTTTCAATCTCTTGGCTTGTCTCAAGGTG 60
Qy 61 CTTTAGGCTCGCTGCTTTCTCATGGGCATCCATAGAGTCTTTCAGTCTTTTACCTCGCTC 120
Db 61 CTTTAGGCTCGCTGCTTTCTCATGGGCATCCATAGAGTCTTTCAGTCTTTTACCTCGCTC 120
Qy 121 CGAGGCGATCTTCACTGTATAACCTTCAGGCGCCGAATCATACCACTACTCTTTACAGGCA 180
Db 121 CGAGGCGATCTTCACTGTATAACCTTCAGGCGCCGAATCATACCACTACTCTTTACAGGCA 180
Qy 181 ATTTTTTTAGACATCTCTCAGCTCGTACAGGTGAAGAGCATGCGAAGTACAGAGAAAT 240
Db 181 ATTTTTTTAGACATCTCTCAGCTCGTACAGGTGAAGAGCATGCGAAGTACAGAGAAAT 240
Qy 241 ACGAAGCACCTCCGGTTTCTGGATCGCTGGAGCACCCGCTCTTGAATCGACCGATC 300
Db 241 ACGAAGCACCTCCGGTTTCTGGATCGCTGGAGCACCCGCTCTTGAATCGACCGATC 300
Qy 301 CGAAGTCTTCAACCATGTGATGAAGAGCCTACGACTATCCGAACCTGTGTGGCGG 360
Db 301 CGAAGTCTTCAACCATGTGATGAAGAGCCTACGACTATCCGAACCTGTGTGGCGG 360
Qy 361 CTCAGTGTCTCAGAAATTCGTAACCGAGATGTTGTTTACGGCGGAAGGTGAAGTCTATA 420
Db 361 CTCAGTGTCTCAGAAATTCGTAACCGAGATGTTGTTTACGGCGGAAGGTGAAGTCTATA 420
Qy 421 ACGGACATCGAAGGATCATGATCCCTCTCTGTCGCTCAGGCGGTAAGTCGATGTC 480
Db 421 ACGGACATCGAAGGATCATGATCCCTCTCTGTCGCTCAGGCGGTAAGTCGATGTC 480
Qy 481 CAATTTTCTTAGAAAAAGTATGGAACCTTGTCGACAAGATGATGAGGAGTGGGCTGAGA 540
Db 481 CAATTTTCTTAGAAAAAGTATGGAACCTTGTCGACAAGATGATGAGGAGTGGGCTGAGA 540
Qy 541 AGGATATGCGCTGGGAGAGTCGCGCGGTGAAAAGAGGCAACCAAGACTCGAGACCGAG 600
Db 541 AGGATATGCGCTGGGAGAGTCGCGCGGTGAAAAGAGGCAACCAAGACTCGAGACCGAG 600
Qy 601 GAGTCGATGTAAAGGATGGGTTCGCTCAGACTACTCTGGACGCTCATGGCTCTTTCAGGAT 660
Db 601 GAGTCGATGTAAAGGATGGGTTCGCTCAGACTACTCTGGACGCTCATGGCTCTTTCAGGAT 660
Qy 661 TTGCACTAAGAGCGACTCGCTTCAGAAACAAGACCAATGAGTCTATGTCGCTTTTTCG 720
Db 661 TTGCACTAAGAGCGACTCGCTTCAGAAACAAGACCAATGAGTCTATGTCGCTTTTTCG 720
Qy 721 GACTTACCGATGGGTTTGCTCTACTCTGGACTCGTTTCAAGGCTATCATGTGGGATTTTG 780
Db 721 GACTTACCGATGGGTTTGCTCTACTCTGGACTCGTTTCAAGGCTATCATGTGGGATTTTG 780
Qy 781 TACCTTACTTCGGAACCTATGAAAACGGAGACATGAGATACCTTTGACTCAAGGATTAGCG 840
Db 781 TACCTTACTTCGGAACCTATGAAAACGGAGACATGAGATACCTTTTACTCAAGGATTAGCG 840
Qy 841 TTTCCCGACGAGTTGGGATTCGAGCTTATGAGGACAAAAGAGCAGGCGCTGTTGGCTCAG 900
Db 841 TTTCCCGACGAGTTGGGATTCGAGCTTATGAGGACAAAAGAGCAGGCGCTGTTGGCTCAG 900
Qy 901 CTTTCGATCAGGCTGTGTGATAAAAAGGATGTTCAAGGTTCGGGATATCTTAAGTCTCTTAG 960
Db 901 CTTTCGATCAGGCTGTGTGATAAAAAGGATGTTCAAGGTTCGGGATATCTTAAGTCTCTTAG 960

```

Qy	961	TCAGAGCAACATCGCGCCAACTGCCTGAACTCTCAAAAGCTCTCGATGAGGAGGTAC	1020
Db	961	TCAGAGCAACATCGCGCCAACTGCCTGAACTCTCAAAAGCTCTCGATGAGGAGGTAC	1020
Qy	1021	TCGCTCAGATCAGTAAACCTGTTATTTCCTGGGATATGAAACTTCTTCGACAGTCTTGACAT	1080
Db	1021	TCGCTCAGATCAGTAAACCTGTTATTTCCTGGGATATGAAACTTCTTCGACAGTCTTGACAT	1080
Qy	1081	GGATGTTTCAACCGACTCTCAGAAGACAAAGCCGTTCAGGATAAACTTTCGAGAGAATTTT	1140
Db	1081	GGATGTTTCAACCGACTCTCAGAAGACAAAGCCGTTCAGGATAAACTTTCGAGAGAATTTT	1140
Qy	1141	GTFCAGATCGACACGAGATATGCCTTACCTACGTAGACGAACTTTAATGCGTTCGCTTTATCTCGAAG	1200
Db	1141	GTFCAGATCGACACGAGATATGCCTTACCTACGTAGACGAACTTTAATGCGTTCGCTTTATCTCGAAG	1200
Qy	1201	CGTTCGTTTAAGGAGTCTCTTCGCTAGACCCCTCTAGTCOGTATGCTAAACCGTGAATGCT	1260
Db	1201	CGTTCGTTTAAGGAGTCTCTTCGCTAGACCCCTCTAGTCOGTATGCTAAACCGTGAATGCT	1260
Qy	1261	TAAAGGATGAAGACTTCATCCCACTTCCCGAGCCCTGTCAATTGGTCGAGATGGGTTCGCTCA	1320
Db	1261	TAAAGGATGAAGACTTCATCCCACTTCCCGAGCCCTGTCAATTGGTCGAGATGGGTTCGCTCA	1320
Qy	1321	TCACGAGGTCGGATCACAAGAAAGAACGATGGTCAATGCTTCGTTGTTTCAAACATCAATC	1380
Db	1321	TCACGAGGTCGGATCACAAGAAAGAACGATGGTCAATGCTTCGTTGTTTCAAACATCAATC	1380
Qy	1381	GTTCCAAAGTTCATTTATGGAGAGAGATGCAGAAAGAAATTCAGACCCGAGAGGTGGCTTGAGG	1440
Db	1381	GTTCCAAAGTTCATTTATGGAGAGAGATGCAGAAAGAAATTCAGACCCGAGAGGTGGCTTGAGG	1440
Qy	1441	ACGTAAACAGACTCGCTCAACAGTATTGGAAGACCCCTATGGAACAAGCGAGACTTTATCT	1500
Db	1441	ACGTAAACAGACTCGCTCAACAGTATTGGAAGACCCCTATGGAACAAGCGAGACTTTATCT	1500
Qy	1501	CTGACCACAGAGCTTGCTTTGGTTGGCGAATTGCTGTCCCGAGATGAAGGCCCTTCTTGT	1560
Db	1501	CTGACCACAGAGCTTGCTTTGGTTGGCGAATTGCTGTCCCGAGATGAAGGCCCTTCTTGT	1560
Qy	1561	TTGTCACTCTCCGTCGGGTCAGGTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACA	1620
Db	1561	TTGTCACTCTCCGTCGGGTCAGGTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACA	1620
Qy	1621	TCACCTTGATCATTTCCGTCCTCGAATCGTTGGTAGAGAGAGGAGGGGTACCCAGATGC	1680
Db	1621	TCACCTTGATCATTTCCGTCCTCGAATCGTTGGTAGAGAGAGGAGGGGTACCCAGATGC	1680
Qy	1681	GTTCGAGGTCAGCCGGTCCGAATGAGTTGATTCCTCATATGTTTAAGAGAAGTTCATAT	1740
Db	1681	GTTCGAGGTCAGCCGGTCCGAATGAGTTGATTCCTCATATGTTTAAGAGAAGTTCATAT	1740
Qy	1741	CTGGAATGCTGACTAGGACAATGCTTCTTTGTTATCGATTTGTTTCTCATACCCGGGC	1800
Db	1741	CTGGAATGCTGACTAGGACAATGCTTCTTTGTTATCGATTTGTTTCTCATACCCGGGC	1800
Qy	1801	AGGCGCTATGACTTCTACGTCGCTATATCGCTCTGGACTCTCTTCTTACCCCTATAT	1860
Db	1801	AGGCGCTATGACTTCTACGTCGCTATATCGCTCTGGACTCTCTTCTTACCCCTATAT	1860
Qy	1861	TATTCCTATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCCGGCT	1920
Db	1861	TATTCCTATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCCGGCT	1920
Qy	1921	CGTCGCCGAATTC	1932
Db	1921	CGTCGCCGAATTC	1932

RESULT 2					
AX034665					
LOCUS	AX034665	1932 bp	DNA	linear	PAT 15-NOV-2000
DEFINITION	Sequence 2 from Patent EP1035206.				

AX034665
AX034665.1 GI:11190638
Xanthophyllomyces dendrorhous (anamorph: Phaffia rhodozyma)
Xanthophyllomyces dendrorhous
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Cystofilobasidiales;
Cystofilobasidiaceae; Xanthophyllomyces.
1.
Hoshino, T., Ojima, K. and Setoguchi, Y.
Astaxanthin synthetase
Patent: EP 1035206-A 2 13-SEP-2000;
HOFFMANN LA ROCHE (CH)
Location/Qualifiers
1. .1932
/organism="Xanthophyllomyces dendrorhous"
/mol_type="unassigned DNA"
/db_xref="taxon:5421"
14. .1891
33. .1706
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC16261.1"
/db_xref="GI:11190639"
/db_xref="REVIEW:EMBL: CAC16261"
/translation="MFILVLLTGALGLAASFWSIAFPLSLYLAPRRSLVNLQPNHT
NYPTGNFLDILSARTGEBEHAKRYKGYSTLRFAGTAGAPVLNSDTPKVFHVMKEAYD
YKPGMAARVLRIATGDGVVTAEGAEHRRHRIIMPSLSAQAVKSVPIPLEKGMELV
DKMDEADKAVGESAGEKATLEETEGVDVDMVGRATLDVMALAGDPYKSDSLQ
NKTNELYAVFVLGTGPAPLDSKRAIMWDFVPRMKRHEIPLTQGLAVRRVGI
ELMEKKQAVLSADQAVDKDVOGRDILSLVLANIAPESQKLSDEEVLQIS
NLIFAGTSSTVLSQWHRLESDRAVDKRELCQIDTMDPLDELNALPYLEAFV
KESLRDPPSPVANRECUKEDFIPLBEVIVGRDGSVNEVRITKNTMVMPLFNINR
SKFTYGEDAEERPERWLEDVDSLSNEAPYQASFIQSPRACFGRRFAVEMKAF
LFTVLRVQFEPIISHPEYEHITLIISPRIVGREKEGYQMRLOKVE"
1871
polya_site
ORIGIN
Query Match 100.0%; Score 1932; DB 6; Length 1932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAATTCGGACGAGGCGACCTACTTCTCCATATGTTTCATCTTGTCTTGTCTTCACAGTG 60
Db 1 GAATTCGGACGAGGCGACCTACTTCTCCATATGTTTCATCTTGTCTTGTCTTCACAGTG 60
Qy 61 CTTTAGCGCTGGCTGTTTCTCATGGGCATCCATAGCGTTCATAGTCTTACCTCGCTC 120
Db 61 CTTTAGCGCTGGCTGTTTCTCATGGGCATCCATAGCGTTCATAGTCTTACCTCGCTC 120
Qy 121 CGAGGCGATCTTCACTGTATACCTTCAGGCGCCGAATACCACTACTTTACAGGCA 180
Db 121 CGAGGCGATCTTCACTGTATACCTTCAGGCGCCGAATACCACTACTTTACAGGCA 180
Qy 181 ATTTTATAGACATCTCTCAGCTCGTACAGGTGAAGAGCATCGAAGTACAGAGAAAT 240
Db 181 ATTTTATAGACATCTCTCAGCTCGTACAGGTGAAGAGCATCGAAGTACAGAGAAAT 240
Qy 241 ACAGAGCACCTTCGGTTCGTGGATCGCTGGAGCACCGCTTGTGAACCTGACCGATC 300
Db 241 ACAGAGCACCTTCGGTTCGTGGATCGCTGGAGCACCGCTTGTGAACCTGACCGATC 300
Qy 301 CGAAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCCGAAACCTGGTATGGCG 360
Db 301 CGAAAGTCTTCAACCATGTGATGAAGAAGCCTACGACTATCCGAAACCTGGTATGGCG 360
Qy 361 CTGAGTGCTCAGAAATGCTACCGAGATGTTGTTTACGCGGGAAGTGAGCTCAT 420
Db 361 CTGAGTGCTCAGAAATGCTACCGAGATGTTGTTTACGCGGGAAGTGAGCTCAT 420
Qy 421 AGCGACATCGAAGATCATGATCCCTCTCTGTCCGCTCAGCGCGTTAAGTCGATGTC 480
Db 421 AGCGACATCGAAGATCATGATCCCTCTCTGTCCGCTCAGCGCGTTAAGTCGATGTC 480

Qy 481 CAATTTTCTTAAAAAGGTATGAACTTGTGCAAGATGATCGAGGATCGCGCTGAGA 540
Db 481 CAATTTTCTTAAAAAGGTATGAACTTGTGCAAGATGATCGAGGATCGCGCTGAGA 540
Qy 541 AGGATATGCGCGTGGAGAGTGGCCCGTGAAGAAGGCAACGAGACTCGAGACCGAAG 600
Db 541 AGGATATGCGCGTGGAGAGTGGCCCGTGAAGAAGGCAACGAGACTCGAGACCGAAG 600
Qy 601 GAGTCGATGTAAGAGATTGGGTGCGTCCGAGCTACTCTGGACGTCATGGCTTCTGAGGAT 660
Db 601 GAGTCGATGTAAGAGATTGGGTGCGTCCGAGCTACTCTGGACGTCATGGCTTCTGAGGAT 660
Qy 661 TTGACTATPAGAGCGACTCGCTCCAGAACAGCAATAGAGCTCTATGTGCTTTTGTGCG 720
Db 661 TTGACTATPAGAGCGACTCGCTCCAGAACAGCAATAGAGCTCTATGTGCTTTTGTGCG 720
Qy 721 GACTTACCGATGGGTGCTCTTACCTTGGACTCGTTCAGGCTATCATGTGGGATTTTG 780
Db 721 GACTTACCGATGGGTGCTCTTACCTTGGACTCGTTCAGGCTATCATGTGGGATTTTG 780
Qy 781 TACCTTACTTCCGAACATATGAACCGAGACATGAGATACCTTTGACTCAAGGATTAGCAG 840
Db 781 TACCTTACTTCCGAACATATGAACCGAGACATGAGATACCTTTGACTCAAGGATTAGCAG 840
Qy 841 TTTCCGACGAGTTGGGATCGAGCTTATGAGCAAAAGAGCAGCGCGCTGCTTGGCTCAG 900
Db 841 TTTCCGACGAGTTGGGATCGAGCTTATGAGCAAAAGAGCAGCGCGCTGCTTGGCTCAG 900
Qy 901 CTTCCGATCAGGCTGTTTCATAAAAAGGATGTTCAAGGTCGGGATATCCTTAAGTCTCCTAG 960
Db 901 CTTCCGATCAGGCTGTTTCATAAAAAGGATGTTCAAGGTCGGGATATCCTTAAGTCTCCTAG 960
Qy 961 TGAGAGCAAAATATCGCCCGCAACCTGCTGAATCTCAAAAGCTGTCGAGTGAAGGATAC 1020
Db 961 TGAGAGCAAAATATCGCCCGCAACCTGCTGAATCTCAAAAGCTGTCGAGTGAAGGATAC 1020
Qy 1021 TCGCTCAGATCAGTAACCTGTTTATTTGCTGGATATGAAACTTCTTCACAGTCTTGCAT 1080
Db 1021 TCGCTCAGATCAGTAACCTGTTTATTTGCTGGATATGAAACTTCTTCACAGTCTTGCAT 1080
Qy 1081 GGATGTTTACCGACTCTCAGAGCAAAAGCGGTTCCAGGATAAACTTCGAGAGAAATTT 1140
Db 1081 GGATGTTTACCGACTCTCAGAGCAAAAGCGGTTCCAGGATAAACTTCGAGAGAAATTT 1140
Qy 1141 GTGAGTCGACACGATATGCTACGCTAGACGAACTTAATGCGTTCCTTATCTCGAAG 1200
Db 1141 GTGAGTCGACACGATATGCTACGCTAGACGAACTTAATGCGTTCCTTATCTCGAAG 1200
Qy 1201 CGTTTGTAAAGGAGTCTCTTCTGCTAGACCTCTCTAGTCCGATGCTAAACCGTGAATGCT 1260
Db 1201 CGTTTGTAAAGGAGTCTCTTCTGCTAGACCTCTCTAGTCCGATGCTAAACCGTGAATGCT 1260
Qy 1261 TAAAGATGAAGACTTCTCCACTTGGCGAGCTGCTATGTTGGTGGAGTGGGTGCGTCA 1320
Db 1261 TAAAGATGAAGACTTCTCCACTTGGCGAGCTGCTATGTTGGTGGAGTGGGTGCGTCA 1320
Qy 1321 TCAACGAGTCCGGATCAGAAACGATGCTGCTGCTTCCGTTGTTCAACATCAATC 1380
Db 1321 TCAACGAGTCCGGATCAGAAACGATGCTGCTGCTTCCGTTGTTCAACATCAATC 1380
Qy 1381 GTTCAAGTTCATTATGGAAGATGCAAGAAATTCAGACCGGAGAGTGGCTTGAGG 1440
Db 1381 GTTCAAGTTCATTATGGAAGATGCAAGAAATTCAGACCGGAGAGTGGCTTGAGG 1440
Qy 1441 AGTAAACAGCTCGCTCAACAGTATGGAACACCTATGGAACACCGAGCGAGCTTTATCT 1500
Db 1441 AGTAAACAGCTCGCTCAACAGTATGGAACACCTATGGAACACCGAGCGAGCTTTATCT 1500
Qy 1501 CTGGACCCAGAGCTTCTGTTGGCGAATTTGCTGTCCCGAGATGAAGCCCTTTTGT 1560
Db 1501 CTGGACCCAGAGCTTCTGTTGGCGAATTTGCTGTCCCGAGATGAAGCCCTTTTGT 1560

	Qy	1561	TTGTGCACTCTCGTCGGTCCAGTTTCGAGGCCCATCATCTCTCATCCAGAGTACGAGCACA	1620
	Dd	1561	TTGTGCACTCTCGTCGGTCCAGTTTCGAGGCCCATCATCTCTCATCCAGAGTACGAGCACA	1620
	Qy	1621	TCACTTTGATCATTTCCCGTCTCGAATCGTTGGTAGAGAAGAGGGGTACCAGATGC	1680
	Dd	1621	TCACTTTGATCATTTCCCGTCTCGAATCGTTGGTAGAGAAGAGGGGTACCAGATGC	1680
	Qy	1681	GTTTGCAAGTCAAAGCCGGTCCAATGAGTTGAATTCATTGTAAGAGAAATTTCTATAT	1740
	Dd	1681	GTTTGCAAGTCAAAGCCGGTCCAATGAGTTGAATTCATTGTAAGAGAAATTTCTATAT	1740
	Qy	1741	CTGAGAACTGTGACTAGGACAAATGCCTCTTTGTATCGAATTTGTTTCTCATACCCGGGC	1800
	Dd	1741	CTGAGAACTGTGACTAGGACAAATGCCTCTTTGTATCGAATTTGTTTCTCATACCCGGGC	1800
	Qy	1801	AGCGCGTATGACTTCTACGTCGTATCTGTCGCTCTGGACTCTCTTTTACCTTATATAT	1860
	Dd	1801	AGCGCGTATGACTTCTACGTCGTATCTGTCGCTCTGGACTCTCTTTTACCTTATATAT	1860
	Qy	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCCGCT	1920
	Dd	1861	TATTCATCCGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAGCCGCT	1920
	Qy	1921	CGTGGCGAATTC 1932	
	Dd	1921	CGTGGCGAATTC 1932	
		RESULT 3		
		AR203450		
		LOCUS	AR203450	linear PAT 20-JUN-2002
		DEFINITION	Sequence 4 from patent US 6365386.	
		ACCESSION	AR203450	
		VERSION	AR203450.1 GI:21499842	
		KEYWORDS	Unknown.	
		SOURCE	Unknown.	
		ORGANISM	Unclassified.	
		REFERENCE	1 (bases 1 to 3969)	
		AUTHORS	Hoshino,T., Ojima,K. and Setoguchi,Y.	
		TITLE	Astaxanthin synthase	
		JOURNAL	Patent: US 6365386-A 4 02-APR-2002;	
		FEATURES	Location/Qualifiers	
		source	1..3969	
		/organism="unknown"		
		/mol_type="unassigned DNA"		
		ORIGIN		
		Query Match	18.3%; Score 353.2; DB 6; Length 3969;	
		Best Local Similarity	99.2%; Pred. No. 3,7e-96;	
		Matches 355; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
	Qy	1514	TTGCTTTGGTTGGCGAATTTGCTGTGCGCGAGATGAAGGCCCTCTTGTGTTGCATCTCCG	1573
	Dd	3508	TCGCAGTGGTTGGCGAATTTGCTGTGCGCGAGATGAAGGCCCTCTTGTGTTGCATCTCCG	3567
	Qy	1574	TCGGGTCCAGTTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACATCACCTTGATCAT	1633
	Dd	3568	TCGGGTCCAGTTTCGAGCCCATCATCTCTCATCCAGAGTACGAGCACATCACCTTGATCAT	3627
	Qy	1634	TTCCCGTCTCGAATCGTTGGTTAGAGAAAGAGGGGTACCGAGATCGGTTTGCAGGTCAA	1693
	Dd	3628	TTCCCGTCTCGAATCGTTGGTTAGAGAAAGAGGGGTACCGAGATCGGTTTGCAGGTCAA	3687
	Qy	1694	GCGGTCGAAATGAGTTGATTTCTCATATGTTTAAGAGAAATGTTCTATATCTGAGAAATGTGTG	1753
	Dd	3688	GCGGTCGAAATGAGTTGATTTCTCATATGTTTAAGAGAAATGTTCTATATCTGAGAAATGTGTG	3747
	Qy	1754	ACTAGAACAAATGCCCTCTTTGTATCGAATTTGTTTCTCATACCCGGCGAGCGCTATGACT	1813
	Dd	3748	ACTAGAACAAATGCCCTCTTTGTATCGAATTTGTTTCTCATACCCGGCGAGCGCTATGACT	3807

Qy	1814	TCTAGTCGCTATCGTCGGCTCTGGACTCTCTCTTACCCCTATATATATTTCCATCCG	1871
Db	3808	TCTAGTCGCTATCGTCGGCTCTGGACTCTCTCTTACCCCTATATATTTCCATCCG	3865
RESULT 4			
AX034666			
LOCUS	AX034666	3969 bp	DNA
DEFINITION	Sequence 3 from Patent EP1035206.		linear
ACCESSION	AX034666		
VERSION	AX034666.1	GI:11190640	
KEYWORDS			
SOURCE	Xanthophyllomyces dendrorhous (anamorph: Phaffia rhodozyma)		
ORGANISM	Xanthophyllomyces dendrorhous		
REFERENCE	Eukaryota; Fungi; Basidiomycota; Hymenomycetes;		
AUTHORS	Heterobasidiomycetes; Tremellomycetidae; Cystofilobasidiales;		
TITLE	Cystofilobasidiaceae; Xanthophyllomyces.		
JOURNAL	1		
FEATURES			
source	Hoshino, T., Ojima, K. and Setoguchi, Y.		
5'UTR	Patent: EP 1035206-A 3 13-SEP-2000;		
exon	KOFFMANN LA ROCHE (CH)		
intron	Location/Qualifiers		
exon	1..3969		
intron	/organism="Xanthophyllomyces dendrorhous"		
exon	/mol_type="unassigned DNA"		
intron	/db_xref="taxon:5421"		
exon	517..518		
intron	/note="EXPERIMENTAL"		
exon	535..652		
intron	653..734		
exon	735..783		
intron	784..898		
exon	899..1015		
intron	1016..1087		
exon	1088..1179		
intron	1180..1302		
exon	1303..1517		
intron	1518..1600		
exon	1601..1634		
intron	1635..1723		
exon	1724..1866		
intron	1867..1939		
exon	1940..1999		
intron	2000..2081		
exon	2082..2181		
intron	2182..2257		
exon	2258..2354		
intron	2355..2431		
exon	2432..2542		
intron	2543..2618		
exon	2619..2652		
intron	2653..2742		
exon	2743..2814		
intron	2815..2962		
exon	2963..3050		
intron	3051..3113		
exon	3114..3171		
intron	3172..3247		
exon	3248..3321		
intron	3322..3398		
exon	3399..3423		
intron	3424..3513		
exon	3514..3700		
polyA_site	3865..3866		
ORIGIN	/note="EXPERIMENTAL"		
Query Match	18.3%;	Score 353.2;	DB 6; Length 3969;
Best Local Similarity	99.2%;	Pred No. 3.7e-96;	
Matches 355;	Conservative	0; Mismatches	3; Indels 0; Gaps 0;
Qy	1514	TTGCTTTGGTGGCGAATTGCTGTCGCCGAGATGAAGGCGCTTCTTGTTGTACTTCG	1573


```

QY 1117 AGGATAAACTTCGAGAGAAATT-----TGTGATGTCGACGAGATATGCTTACGCTAG 1170
Db 1025 AGCAGAGATTCATGAGAGATAGATTCAATTTCTCCCTGCAAGGCCAGTCCCTACTTATG 1084
QY 1171 ACAGAACTTAATGCGTTCCTTATCTCGAAGGTTTGTAAAGGAGTCTCTTCGCTAGACC 1230
Db 1085 ACATTCGATCGCATGATGAATATCTTGTATGTCATTCAGGAGACCCCTCAGGTTGATC 1144
QY 1231 CTCCT 1235
Db 1145 CCCCT 1149

RESULT 6
LOCUS AX655393 2000 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 5263 from Patent WO03000898.
ACCESSION AX655393
VERSION AX655393.1 GI:29158207
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE 1
AUTHORS Chang H.S., Chen.W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
TITLE Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
JOURNAL Plant genes involved in defense against pathogens
Patent: WO 03000898-A 5263 03-JAN-2003;
Syngenta Participation AG (CH)
FEATURES
source
1..2000
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Query Match 2.7%; Score 51.8; DB 6; Length 2000;
Best Local Similarity 9.8%; Pred. No. 0.0011;
Matches 72; Conservative 343; Mismatches 310; Indels 10; Gaps 3;

QY 421 ACCGATCGAAGGATCATGATCCCTCTCTGCGCTCAGCGCGTTAACTCATGCTCC 480
Db 48 RCGSCGCKMTRKSKWYSASSAGRTGSKWSGSGYSGMKKRYKRGRGRGR 107
QY 481 CAATTTCTTAGAAAAAGTATGGAACCTGTGCAAGATGATGAGGATCGCGCTGAGA 540
Db 108 MRSRMRWGRYRCARSGRMAGGSRMVGKSRMSYMMCYARGCGSKRKS--KGS 165
QY 541 AGGATATGCCGTGGAGATGCGCGCGTGAAGAAGAGGCAACAGACTCGACACGAG 600
Db 166 WGTCTRGARGGSGWSGAKYKSGMSXRMWSSCGRSRGSAYRYGTGRKYGTYK 225
QY 601 GAGTCATGTAAGGATGGGTGCGTTCGAGCTACTCTGGAGCTCATGGCTCTTGCAGGAT 660
Db 226 KMTYSASECMRAYMTTSYACSSYTWCRSKRSMMWKMVKRWSRYGWSYKWM 285
QY 661 TTGACTATAAGAGCGACTGCTGCCAGACAAGACCAATGACTCTATGTCGCTTTTCTCG 720
Db 286 MCTAYKYSYRWCYMYRGGRGATRYWGRGYMSRMAMTKYMYRYGYKMKRGMWAG 345
QY 721 GACTTACGATGGGTTGCTCTCTACTCTTGGACTGCTTCAAGGCTCATATGCGGATTTG 780
Db 346 RMMRSMCRWSKACYTMRWMTMRWMTRRRWAKKSRSTRSRKRWKMKRKYKMRYSR 405
QY 781 TACCTTACTCCGAACATATGAACGAGACATGAGA-----TACCTTTGACTCAAGAT 834
Db 406 MRSCKRMRWKRCSGRAWMGRCGTCRMKSYGMWKMWSKMRKMSKYSKMSMYRWK 465
QY 835 TAGCAGTTTCCGACGAGCTTGGATCGAGCTTATGAGCAAAAGACGACGCGCTGCTG 894
Db 466 KCSRTTMGKTRGGMMGTGRCRYKRSKMRKCRRRRGRMGRMYMRWRYMSARYTMR 525

```

```

QY 895 GCTCAGCTTCCGATCAGGCTGTGTGATAAAAAAGAGTGTTCAAGGTCGGGATATCCTAAGTC 954
Db 526 YCARKKYSYSAARFARCWYRGKYWAGWMMKRYKRMWYKMMWYKRYKSKSWYCKWS 585
QY 955 TCCTAGTAGAGAGCAACATCGCC--GCCAACCTGCTGAATCTCAAAAGCTGTCCGATGA 1012
Db 586 YIASCKMSARKAGAKWCKESKMSANSKSMRSRCKRCKASKRESSAKRYAMWGMGTSGSRM 645
QY 1013 GGAGGTACTCGTCAGATCAGTAACCTGTATTGCTGGATATGAACCTTCTTCGACAGT 1072
Db 646 SRWKSCTYWRKWSMKSTCTWYNSKYTYAKYSGYWRYYRACWCMYWRWYRYRYKSY 705
QY 1073 CTTGACATGATGTTTCAAGGCTCTCAAGACACAAAGCCGTCAGGATAAACTTCGAGA 1132
Db 706 MTYAWVTSSTRMAMTGKYSGRYWTWYKCKSKWYSMMYYWWSWAKTWKWRRYA 765
QY 1133 AGAAATTTGTCAGAT 1147
Db 766 TRMMWYRYSMKWY 780

RESULT 7
LOCUS HSA563377 1062 bp mRNA linear PRI 15-MAY-2003
DEFINITION Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION AJ563377
VERSION AJ563377.1 GI:30840240
KEYWORDS CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Finta,C. and Zaphiropoulos,P.G.
TITLE Intergenic mRNA molecules resulting from trans-splicing
JOURNAL J. Biol. Chem. 277 (8), 5882-5890 (2002)
MEDLINE 21839017
PUBMED 11726664
REFERENCE 2 (bases 1 to 1062)
AUTHORS Zaphiropoulos,P.G.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska
Institute, Novum, 141 57 Huddinge, SWEDEN
FEATURES
source
1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
1..1062
/genes="CYP3A43/CYP3A4"
/genes="CYP3A43/CYP3A4"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAD91345.1"
/db_xref="GI:30840241"
/translation="MDLIENFAMETWLVATSLVLLYIVFGAYSMVITSTSGVNIID
SLNPPQFFVENTKLRDFDLDPFLSTVFPFLIPLELVNICVFREVTNFKRS
VKRKSRLDEDTQKRVDFQLMDSQSKTESKHALSDLELVASQIIFIAGIETT
SSVLSTFMYELATHPDVQOKLQDEIDAVLPNKAPPTDVLQWYLDVWVNETLRF
IDRLERVKCKVEINGMTIPKGVVYVNIYSALHRDPKYTWPEKFLPERFSKKNKD
IDPYVTFPGSGPRNCIGMRFALNMKALIRVLQNFSPKPKETQIPLKLSIGLLQ
PEKPVVLKVSRRDGTYSGA"

ORIGIN
Query Match 2.5%; Score 48.4; DB 9; Length 1062;
Best Local Similarity 50.7%; Pred. No. 0.01;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAAGCTGTCGAGAGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTGCTGGATATG 1056

```

```
Db 413 AAGCTCTGCGATCTGGAGCTCGTGGCCCAATCAATTATCTTTATTTTGGCTGCTATG 472
QY 1057 AAACCTTTCTCGACAGCTTTGACATGGATGTTTTCACCGACTCTCAGAAAGACAAAGCCGTTTC 1116
Db 473 AAACACGAGCAGTGTCTCTCTCTCAATATGATGAACTGGCCACTCACCCTGATGCTCC 532
QY 1117 AGGATTAACCTTCGAGAGAAATTTTCAGATCGACACGGATATCCCTACGTCAGTAGACGAAC 1176
Db 533 AGCAGAAATCTCAGGAGGAAATGTATGCGAGTTTACCCTAAGGACCAACCCACCTATG 592
QY 1177 TTAATCGCTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGACC 1230
Db 593 ATACTGTGTACAGATGGAGTATCTTGACATGGTGTGAATGAAGACGCTCAGATTATTC 652
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGATGAAGACTTCA 1278
Db 653 CAATTGCTATGAGACTTGAGAGGGTCTGCAGAAAAGATGTTGAGATCA 700

RESULT 8
LOCUS HSA563376 1365 bp mRNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens partial mRNA for cytochrome P450 (CYP3A43 and CYP3A4
genes).
ACCESSION AJ563376
VERSION AJ563376.1 GI:30840238
KEYWORDS CYP3A4 gene; CYP3A43 gene; cytochrome P450.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Finta,C. and Zaphiropoulos,P.G.
TITLE Intergenic mRNA molecules resulting from trans-splicing
JOURNAL J. Biol. Chem. 277 (8), 5882-5890 (2002)
MEDLINE 2183017
PubMed 11728664
REFERENCE 2 (bases 1 to 1365)
AUTHORS Zaphiropoulos,P.G.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska
Institute, Novum, 141 57 Huddinge, SWEDEN

FEATURES
source
1..1365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
1 71
/genes="CYP3A43"
1..>71
/genes="CYP3A43"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAD91644.1"
/db_xref="GI:31043862"
/translation="MDLIENFAMETWLVATSLVLLY"
72..1365
/genes="CYP3A4"
<72..1365
/genes="CYP3A4"
/codon_start=2
/product="cytochrome P450"
/protein_id="CAD91645.1"
/db_xref="GI:31043863"

gene
1..1365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
1 71
/genes="CYP3A43"
1..>71
/genes="CYP3A43"
/codon_start=1
/product="cytochrome P450"
/protein_id="CAD91644.1"
/db_xref="GI:31043862"
/translation="MDLIENFAMETWLVATSLVLLY"
72..1365
/genes="CYP3A4"
<72..1365
/genes="CYP3A4"
/codon_start=2
/product="cytochrome P450"
/protein_id="CAD91645.1"
/db_xref="GI:31043863"

gene
1..1365
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

CDS
1..1365
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

gene
1..1512
/organism="Homo sapiens"
/tissue_type="liver"
1..1512
/product="cytochrome P4503A4"

FEATURES
source
1..1512
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTCTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056
Db 716 AAGCTCTCTCGATCTGGAGCTCGTGGCCCAATCAATTATCTTTATTTTGGCTGCTATG 775
QY 1057 AAACCTTTCTCGACAGCTTTGACATGGATGTTTTCACCGACTCTCAGAAAGACAAAGCCGTTTC 1116
Db 776 AAACACGAGCAGTGTCTCTCTCTCAATATGATGAACTGGCCACTCACCCTGATGCTCC 835
QY 1117 AGGATTAACCTTCGAGAGAAATTTTCAGATCGACACGGATATCCCTACGTCAGTAGACGAAC 1176
Db 836 AGCAGAAATCTCAGGAGGAAATGTATGCGAGTTTACCCTAAGGACCAACCCACCTATG 895
QY 1177 TTAATCGCTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGCTAGACC 1230
Db 896 ATACTGTGTACAGATGGAGTATCTTGACATGGTGTGAATGAAGACGCTCAGATTATTC 955
QY 1231 CTCCTAGTCCGTATGCTAACCGTGAATGCTTAAAGATGAAGACTTCA 1278
Db 956 CAATTGCTATGAGACTTGAGAGGGTCTGCAGAAAAGATGTTGAGATCA 1003

RESULT 9
LOCUS E10636 1512 bp RNA linear PAT 29-SEP-1997
DEFINITION Human cDNA encoding cytochrome P4503A4.
ACCESSION E10636
VERSION E10636.1 GI:22027692
KEYWORDS JP 1996027197-A/5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1512)
AUTHORS Funae,Y., Imaoka,S., Matsuki,Y., Hayashi,K. and Yabusaki,Y.
TITLE ANTIBODY RECOGNIZING CYTOCHROME P4502A4 ORIGINATED FROM MAN
JOURNAL Patent: JP 1996027197-A 5 30-JAN-1996;
SUMITOMO CHEM CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996027197-A/5
PD 30-JAN-1996
PF 13-JUL-1994 JP 1994161552
PI FUNAE YOSHIHIKO, IMAOKA SUSUMU, MATSUKI YASUSHI, HAYASHI KOJI,
YABUSAKI YOSHIYASU
PC C07K16/18,C12N15/09//C12N1/19,C12N9/02,G01N33/53,(C12N1/19, PC
C12R1:865);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1512
FT source /organism="Homo sapiens"
FT CDS /tissue_type="liver"
FT CDS 1..1512
FT CDS /product="cytochrome P4503A4"

FEATURES
source
1..1512
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY 997 AAAGCTCTCGATGAGGAGTACTCGCTCAGATCAGTAACCTGTTATTTGCTGGATATG 1056
```

```
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
Db      923  AAACCAAGCAGCAGTGTCTCTCTCTCAATTATGATGAACCTGGCCACTCACCCTGATGTC 982
QY      1117  AGGATAAACTTCGAGAGAATAATTTGTCAGATCGACAGCGATATGCCCTACGCTAGACGAAC 1176
Db      983  ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCACTATG 1042
QY      1177  TTAATGCGTTGC-----CTTATCTCGAAGCGTTTCTTAAGGAGTCTCTTCGCTAGACC 1230
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1103  CAATTGCTATGAGACTTGAGAGGGTCTGCACAAAAGATGTTGAGATCA 1150

RESULT 10
LOCUS   E10855                      1512 bp    RNA        linear    PAT 29-SEP-1997
DEFINITION   CDNA encoding human cytochrome P450.
ACCESSION   E10855
VERSION     E10855.1 GI:22027949
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1512)
AUTHORS    Hayashi,K., Sakaki,T., Yabueki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
TITLE       Method for evaluating safety
JOURNAL     Patent: JP 1996056695-A 4 05-MAR-1996;
            SUMITOMO CHEM CO LTD
COMMENT     OS Homo sapiens (human)
            PN 05-MAR-1996
            PD 05-MAR-1996
            PR 15-JUL-1994 JP 1994164184
            PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
            17-JUN-1994 JP 94P 136053
            PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
            KOICHIRO,
            PI KANEKO HIDEO, NAKATSUKA IWAO
            PC C12Q1/02,C12M1/34,C12Q1/26;
            CC strandedness: Double;
            CC topology: Linear;
            CC hypothetical: No;
            CC anti-sense: No;
            FH Key
            FH Location/Qualifiers
            FT source
            FT 1. .1512
            FT /organism='Homo sapiens'
            FT CDS
            FT 1. .1512
            FT /product='human cytochrome P450 3A4'.
            FT Location/Qualifiers
            FT 1. .1512
            FT /organism='Homo sapiens'
            FT /mol_type='genomic RNA'
            FT /db_xref='taxon:9606'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY      997  AAAAGCTGTCGATGAGGAGTACTCGTTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
Db      923  AAACCAAGCAGCAGTGTCTCTCTCTCAATTATGATGAACCTGGCCACTCACCCTGATGTC 982
QY      1117  AGGATAAACTTCGAGAGAATAATTTGTCAGATCGACAGCGATATGCCCTACGCTAGACGAAC 1176
Db      983  ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCACTATG 1042
QY      1177  TTAATGCGTTGC-----CTTATCTCGAAGCGTTTCTTAAGGAGTCTCTTCGCTAGACC 1230
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1103  CAATTGCTATGAGACTTGAGAGGGTCTGCACAAAAGATGTTGAGATCA 1150

RESULT 12
LOCUS   HSA563375                      1512 bp    mRNA       linear    PRI 15-MAY-2003
DEFINITION   Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION   AJ563375
VERSION     AJ563375.1 GI:30840236
KEYWORDS    CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1512)
AUTHORS    Hayashi,K., Sakaki,T., Yabueki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
TITLE       Method for evaluating safety
JOURNAL     Patent: US 6620593-A 4 16-SEP-2003;
            Location/Qualifiers
            1. .1512
            /organism='unknown'
            /mol_type='genomic DNA'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY      997  AAAAGCTGTCGATGAGGAGTACTCGTTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
Db      923  AAACCAAGCAGCAGTGTCTCTCTCTCAATTATGATGAACCTGGCCACTCACCCTGATGTC 982
QY      1117  AGGATAAACTTCGAGAGAATAATTTGTCAGATCGACAGCGATATGCCCTACGCTAGACGAAC 1176
Db      983  ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCACTATG 1042
QY      1177  TTAATGCGTTGC-----CTTATCTCGAAGCGTTTCTTAAGGAGTCTCTTCGCTAGACC 1230
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1103  CAATTGCTATGAGACTTGAGAGGGTCTGCACAAAAGATGTTGAGATCA 1150

RESULT 12
LOCUS   HSA563375                      1512 bp    mRNA       linear    PRI 15-MAY-2003
DEFINITION   Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION   AJ563375
VERSION     AJ563375.1 GI:30840236
KEYWORDS    CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1512)
AUTHORS    Hayashi,K., Sakaki,T., Yabueki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
TITLE       Method for evaluating safety
JOURNAL     Patent: US 6620593-A 4 16-SEP-2003;
            Location/Qualifiers
            1. .1512
            /organism='unknown'
            /mol_type='genomic DNA'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY      997  AAAAGCTGTCGATGAGGAGTACTCGTTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
```

```
Db      923  AAACCAAGCAGCAGTGTCTCTCTCTCAATTATGATGAACCTGGCCACTCACCCTGATGTC 982
QY      1117  AGGATAAACTTCGAGAGAATAATTTGTCAGATCGACAGCGATATGCCCTACGCTAGACGAAC 1176
Db      983  ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCACTATG 1042
QY      1177  TTAATGCGTTGC-----CTTATCTCGAAGCGTTTCTTAAGGAGTCTCTTCGCTAGACC 1230
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1103  CAATTGCTATGAGACTTGAGAGGGTCTGCACAAAAGATGTTGAGATCA 1150

RESULT 11
LOCUS   AR399323                      1512 bp    DNA        linear    PAT 18-DEC-2003
DEFINITION   Sequence 4 from patent US 6620593.
ACCESSION   AR399323
VERSION     AR399323.1 GI:40141119
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1512)
AUTHORS    Hayashi,K., Sakaki,T., Yabueki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
TITLE       Method for safety evaluation of chemical compound using recombinant
            yeast expressing human cytochrome P450
JOURNAL     Patent: US 6620593-A 4 16-SEP-2003;
            Location/Qualifiers
            1. .1512
            /organism='unknown'
            /mol_type='genomic DNA'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY      997  AAAAGCTGTCGATGAGGAGTACTCGTTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
Db      923  AAACCAAGCAGCAGTGTCTCTCTCTCAATTATGATGAACCTGGCCACTCACCCTGATGTC 982
QY      1117  AGGATAAACTTCGAGAGAATAATTTGTCAGATCGACAGCGATATGCCCTACGCTAGACGAAC 1176
Db      983  ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCACTATG 1042
QY      1177  TTAATGCGTTGC-----CTTATCTCGAAGCGTTTCTTAAGGAGTCTCTTCGCTAGACC 1230
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1103  CAATTGCTATGAGACTTGAGAGGGTCTGCACAAAAGATGTTGAGATCA 1150

RESULT 12
LOCUS   HSA563375                      1512 bp    mRNA       linear    PRI 15-MAY-2003
DEFINITION   Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION   AJ563375
VERSION     AJ563375.1 GI:30840236
KEYWORDS    CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1512)
AUTHORS    Hayashi,K., Sakaki,T., Yabueki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
TITLE       Method for evaluating safety
JOURNAL     Patent: US 6620593-A 4 16-SEP-2003;
            Location/Qualifiers
            1. .1512
            /organism='unknown'
            /mol_type='genomic DNA'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY      997  AAAAGCTGTCGATGAGGAGTACTCGTTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
Db      923  AAACCAAGCAGCAGTGTCTCTCTCTCAATTATGATGAACCTGGCCACTCACCCTGATGTC 982
QY      1117  AGGATAAACTTCGAGAGAATAATTTGTCAGATCGACAGCGATATGCCCTACGCTAGACGAAC 1176
Db      983  ACAGAAATCTGAGGAGAAATTTGATGAGCTTTTACCAATAGGCACACCCACTATG 1042
QY      1177  TTAATGCGTTGC-----CTTATCTCGAAGCGTTTCTTAAGGAGTCTCTTCGCTAGACC 1230
Db      1043  ATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 1102
QY      1231  CTCCTAGTCCGATGCTTACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db      1103  CAATTGCTATGAGACTTGAGAGGGTCTGCACAAAAGATGTTGAGATCA 1150

RESULT 12
LOCUS   HSA563375                      1512 bp    mRNA       linear    PRI 15-MAY-2003
DEFINITION   Homo sapiens mRNA for cytochrome P450 (CYP3A43/CYP3A4 gene).
ACCESSION   AJ563375
VERSION     AJ563375.1 GI:30840236
KEYWORDS    CYP3A43/CYP3A4 gene; cytochrome P450.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1512)
AUTHORS    Hayashi,K., Sakaki,T., Yabueki,Y., Komai,K., Kaneko,H. and
            Nakatsuka,I.
TITLE       Method for evaluating safety
JOURNAL     Patent: US 6620593-A 4 16-SEP-2003;
            Location/Qualifiers
            1. .1512
            /organism='unknown'
            /mol_type='genomic DNA'

ORIGIN
Query Match 2.5%; Score 48.4; DB 6; Length 1512;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY      997  AAAAGCTGTCGATGAGGAGTACTCGTTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
Db      863  AAGCTCTGTCGAGCTGAGCTGGGCCCAATCAATTATCTTTATTTTGTGGCTATG 922
QY      1057  AAACCTCTTCGACAGCTTTGACATGGATGTTTACCGAGCTCTCAGAGAACAAGCGGTTTC 1116
```

1
 Finta, C. and Zaphiropoulos, P.G.
 TITLE
 JOURNAL
 J. Biol. Chem. 277 (8), 5882-5890 (2002)
 MEDLINE
 21839017
 PUBMED
 11726564
 2 (bases 1 to 1512)
 Zaphiropoulos, P.G.
 TITLE
 JOURNAL
 Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska
 Institute, Novum, 141 57 Huddinge, SWEDEN
 FEATURES
 source
 Location/Qualifiers
 1. .1512
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 1. .1512
 /gene="CYP3A43/CYP3A4"
 1. .1512
 /gene="CYP3A43/CYP3A4"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="CAD91343.1"
 /db_xref="GI:30840237"
 /translation="MDLIPFAMETWLVVATSLVLLIYVTHSHGLFKLIGPPTPL
 PFLNLSYHKGFCWDFMECHKYKVMGVDGQVLAITDPMIKTVLVKECVSF
 TNRPFGVPMKSAISARDEWKLRLSLLPTFTSGKLMVPIIAQVGVLRNL
 RREATKPVTLKDFVGNYSMDVITSTFGVNDISLNNPDPEVTKLLRFDELDP
 FFLSITVFFPILEVLNLCVFFREVTNLRKSKVEMKESLEDQKRVDFLOLMI
 IDVQNKTESKALSDLEVAQSILIFIFAGYETTSVLSFIMELAHEDVDQKGEZ
 IDVLPKAPYTVTLQMEYLDVNNETLRLPIAMRLERVCKKQVEINGMFIKGV
 VVMIPSVLHRLDPKWTPEKFLPERSKKNIDPIYIYTFPGSPRCIGNRFALM
 NMLKALIRVLQNFSEKPKCTQIPLKLSGLLQPKPVVLKVESRDGTVSGA"
 ORIGIN
 Query Match 2.5%; Score 48.4; DB 9; Length 1512;
 Best Local Similarity 50.7%; Pred. No. 0.011;
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
 QY 997 AAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAAGTAACTGTTATTTGCTGGATATG 1056
 Db 863 AAGCTCTGTCGATGAGGAGTACTCGCTCAGATCAAGTAACTGTTATTTGCTGGATATG 922
 QY 1057 AAACCTTCTCGACATGATGATGTTTCCAGATCGACAGGATATGCTACGCTAGACGAAC 1176
 Db 923 AAACCTTCTCGACATGATGATGTTTCCAGATCGACAGGATATGCTACGCTAGACGAAC 982
 QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTACGCTAGACGAAC 1176
 Db 983 AGCAGAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTACGCTAGACGAAC 1042
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGTAGACC 1230
 Db 1043 ATACTGTGCTACAGATGAGTATCTTTGATCATGTTGATGAACTGGCCACTCACCCCTGATGTC 1102
 QY 1231 CTCCTAGTCGATGCTACCGTGAATCTTAAAGGATGAGACTTCA 1278
 Db 1103 CAATTGCTATGAGCTTGAGAGGGTCTGCAAAAAGATGTTGATGATCA 1150
 RESULT 13
 AX408939
 LOCUS
 DEFINITION
 Sequence 1586 from Patent WO029103.
 AX408939
 ACCESSION
 AX408939.1 GI:21441644
 VERSION
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Finta, C. and Zaphiropoulos, P.G.
 TITLE
 JOURNAL
 J. Biol. Chem. 277 (8), 5882-5890 (2002)
 MEDLINE
 21839017
 PUBMED
 11726564
 2 (bases 1 to 1512)
 Zaphiropoulos, P.G.
 TITLE
 JOURNAL
 Submitted (14-MAY-2003) Zaphiropoulos P.G., Bioscience, Karolinska
 Institute, Novum, 141 57 Huddinge, SWEDEN
 FEATURES
 source
 Location/Qualifiers
 1. .1512
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 1. .1512
 /gene="CYP3A43/CYP3A4"
 1. .1512
 /gene="CYP3A43/CYP3A4"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="CAD91343.1"
 /db_xref="GI:30840237"
 /translation="MDLIPFAMETWLVVATSLVLLIYVTHSHGLFKLIGPPTPL
 PFLNLSYHKGFCWDFMECHKYKVMGVDGQVLAITDPMIKTVLVKECVSF
 TNRPFGVPMKSAISARDEWKLRLSLLPTFTSGKLMVPIIAQVGVLRNL
 RREATKPVTLKDFVGNYSMDVITSTFGVNDISLNNPDPEVTKLLRFDELDP
 FFLSITVFFPILEVLNLCVFFREVTNLRKSKVEMKESLEDQKRVDFLOLMI
 IDVQNKTESKALSDLEVAQSILIFIFAGYETTSVLSFIMELAHEDVDQKGEZ
 IDVLPKAPYTVTLQMEYLDVNNETLRLPIAMRLERVCKKQVEINGMFIKGV
 VVMIPSVLHRLDPKWTPEKFLPERSKKNIDPIYIYTFPGSPRCIGNRFALM
 NMLKALIRVLQNFSEKPKCTQIPLKLSGLLQPKPVVLKVESRDGTVSGA"
 ORIGIN
 Query Match 2.5%; Score 48.4; DB 9; Length 1512;
 Best Local Similarity 50.7%; Pred. No. 0.011;
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
 QY 997 AAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAAGTAACTGTTATTTGCTGGATATG 1056
 Db 863 AAGCTCTGTCGATGAGGAGTACTCGCTCAGATCAAGTAACTGTTATTTGCTGGATATG 922
 QY 1057 AAACCTTCTCGACATGATGATGTTTCCAGATCGACAGGATATGCTACGCTAGACGAAC 1176
 Db 923 AAACCTTCTCGACATGATGATGTTTCCAGATCGACAGGATATGCTACGCTAGACGAAC 982
 QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTACGCTAGACGAAC 1176
 Db 983 AGCAGAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTACGCTAGACGAAC 1042
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGTAGACC 1230
 Db 1043 ATACTGTGCTACAGATGAGTATCTTTGATCATGTTGATGAACTGGCCACTCACCCCTGATGTC 1102
 QY 1231 CTCCTAGTCGATGCTACCGTGAATCTTAAAGGATGAGACTTCA 1278
 Db 1103 CAATTGCTATGAGCTTGAGAGGGTCTGCAAAAAGATGTTGATGATCA 1150
 RESULT 14
 HUMCYPHLP
 LOCUS
 DEFINITION
 Human liver glucocorticoid-inducible cytochrome P-450 (HLP) mRNA,
 complete cds.
 M13785.1 GI:181353
 VERSION
 KEYWORDS
 SOURCE
 Homo sapiens
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2011)
 REFERENCE
 AUTHORS
 Molowa, D.T., Schuetz, E.G., Wrighton, S.A., Watkins, P.B., Kremers, P.,
 Mender-Picon, G., Parker, G.A. and Guzelian, P.S.
 TITLE
 Complete cDNA sequence of a cytochrome P-450 inducible by
 glucocorticoids in human liver
 JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 83 (14), 5311-5315 (1986)
 MEDLINE
 86259780
 PUBMED
 3450094
 COMMENT
 Original source text: Human (adult) liver, cDNA to mRNA, clones
 Hp[55.12.9].
 Draft entry and sequence in computer-readable form for [1] kindly
 provided by D.T.Molowa, 19-NOV-1986.
 The human genome probably contains multiple genes or pseudogenes
 related to glucocorticoid-inducible cytochrome P-450.
 FEATURES
 source
 Location/Qualifiers
 1. .2011
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 61. .1575
 /note="glucocorticoid-inducible cytochrome P-450"
 /codon_start=1
 /protein_id="AAA35742.1"
 /db_xref="GI:181354"
 CDS
 1

Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
 TITLE
 JOURNAL
 Patent: WO 029103-A 1586 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source
 Location/Qualifiers
 1. .2011
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. D00003"
 ORIGIN
 Query Match 2.5%; Score 48.4; DB 6; Length 2011;
 Best Local Similarity 50.7%; Pred. No. 0.012;
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
 QY 997 AAAAGCTGTCGATGAGGAGTACTCGCTCAGATCAAGTAACTGTTATTTGCTGGATATG 1056
 Db 926 AAGCTCTGTCGATGAGGAGTACTCGCTCAGATCAATTAATCTTTATTTGCTGGCTATG 985
 QY 1057 AAACCTTCTCGACATGATGATGTTTCCAGATCGCTCAGAGCAACAAGCGGCTTC 1116
 Db 986 AAACCTTCTCGACATGATGATGTTTCCAGATCGCTCAGAGCAACAAGCGGCTTC 1045
 QY 1117 AGGATAAATTCGAGAGAAATTTGTCAGATCGACAGGATATGCTACGCTAGACGAAC 1176
 Db 1046 AGCAGAAATTCGAGAGAAATTTGTCAGATCGCTGTTTACCCAAATAGGCACCACTATG 1105
 QY 1177 TTAATGCGTTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGTAGACC 1230
 Db 1106 ATACTGTGCTACAGATGAGTATCTTTGATCATGTTGATGAAAGCGCTCAGATTATTC 1165
 QY 1231 CTCCTAGTCGATGCTACCGTGAATCTTAAAGGATGAGACTTCA 1278
 Db 1166 CAATTGCTATGAGCTTGAGAGGGTCTGCAAAAAGATGTTGATGATCA 1213
 RESULT 14
 HUMCYPHLP
 LOCUS
 DEFINITION
 Human liver glucocorticoid-inducible cytochrome P-450 (HLP) mRNA,
 complete cds.
 M13785.1 GI:181353
 VERSION
 KEYWORDS
 SOURCE
 Homo sapiens
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2011)
 REFERENCE
 AUTHORS
 Molowa, D.T., Schuetz, E.G., Wrighton, S.A., Watkins, P.B., Kremers, P.,
 Mender-Picon, G., Parker, G.A. and Guzelian, P.S.
 TITLE
 Complete cDNA sequence of a cytochrome P-450 inducible by
 glucocorticoids in human liver
 JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 83 (14), 5311-5315 (1986)
 MEDLINE
 86259780
 PUBMED
 3450094
 COMMENT
 Original source text: Human (adult) liver, cDNA to mRNA, clones
 Hp[55.12.9].
 Draft entry and sequence in computer-readable form for [1] kindly
 provided by D.T.Molowa, 19-NOV-1986.
 The human genome probably contains multiple genes or pseudogenes
 related to glucocorticoid-inducible cytochrome P-450.
 FEATURES
 source
 Location/Qualifiers
 1. .2011
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 61. .1575
 /note="glucocorticoid-inducible cytochrome P-450"
 /codon_start=1
 /protein_id="AAA35742.1"
 /db_xref="GI:181354"
 CDS
 1


```
/translation="MALIPDLAVETWLLAVSLVLLVYXGTHSHGLFKKLGIPGTPPL
ZFLGNILSYHKGFCMFMECHCKYKGVYDQGPVLAITDPDMIKLVLVKECVSVF
INREPFPGVGMKSAISAEDEEWKRLRSLSPFTSGKKNWPIIAQKGLVLRNL
RRERETGKVTLDKDFGAYMDVITSSSFGVNDLNNPQDPLVENTKLLRDFLDLP
FFLSITVFPFLIPLEVLNI CVFPRVNTFRKAVRMKESRLEDQKRVDFLOLMI
DSHKNSKTESHSLDLVAQSIIFIFAGYETSSVLSFIMYSLATHPDVQOQLQE
EIDAVLPKAPPTDVLQMEYLDWVNETLAFPIAMRLERVCKDKVEINGMFI PKG
WVMVLPVALHRDPKYWTEPKELPERFSKKNNDIPYITPFGSGPRNCIGMRPAL
NMKMLALIRVLQNSFPKCKETQIPLKLSLGLLQEPKPVLVKVESRDGTVSGA"
```

ORIGIN 170 bp upstream of EcoRI site.

Query Match 2.5%; Score 48.4; DB 9; Length 2011;
Best Local Similarity 50.7%; Pred. No. 0.012;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;

QY	997	AAAGCTGTCGACGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG	1056
DB	926	AAGCTGTGCCATCTGAGCTCGTGGCCCAATCAATATTTATTTTGTGGCTATG	985
QY	1057	AAACTTCTTCGACAGCTTTGACATGATGTTTCCCGACTCTCAGAGACAAAGCCGTTTC	1116
DB	986	AAACCAGGACGAGTGTCTCTCTTCAATATGATGAATCGGCACCTCACCTGATGTC	1045
QY	1117	AGATAAAGCTTCGAGAGAAATTTTCAGATCGACGAGATGCTTACCTAGCAGAAC	1176
DB	1046	AGCAGAACTGCAGAGGAAATTTGATGAGTTTACCAATAGGCACCCACCTATG	1105
QY	1177	TTAATGCGTTCG-----CTTATCTCGAAGCGTGTGTTAAGGAGTCTCTCGTCTAGACC	1230
DB	1106	ATACGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC	1165
QY	1231	CTCCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA	1278
DB	1166	CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA	1213

RESULT 15
LOCUS HUMF450M 2011 bp mRNA linear PRI 13-NOV-1997
DEFINITION Homo sapiens mRNA for cytochrome P-450, complete cds.
ACCESSION D00003 N00003
VERSION D00003.1 GI:219962
KEYWORDS cytochrome P-450.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2011)
Molowa,D.T., Schuetz,E.G., Wrighton,S.A., Watkins,P.B., Kremers,P.,
Mendez-Picon,G., Parker,G.A. and Guzelian,P.S.
Complete cDNA sequence of a cytochrome P-450 inducible by
glucocorticoids in human liver
Proc. Natl. Acad. Sci. U.S.A. 83 (14), 5311-5315 (1986)
JOURNAL MEDLINE 86259780
PUBMED 3460094
COMMENT The liver cytochromes P-450 are supergene family of microsomal
hemoproteins that catalyze the oxidative biotransformation of
numerous endogenous substrates as well as xenobiotics. They used a
monoclonal antibody that recognizes purified human liver cytochrome
P-450 (Hlp) to screen a human liver cDNA library in lambda-gtl.
They isolated and sequenced two overlapping cDNA clones that span
the entire 2011 bases of an mRNA that codes for a protein of 504
amino acids. Hlp mRNA shares more than 70% sequence homology with
related proteins from the rat and rabbit but less than 40% homology
with other published cytochrome P-450 genes. With a
glucocorticoid there was a several fold induction of two mRNA
species hybridizable to Hlp cDNAs as well as an induction of
immunoreactive Hlp protein. They conclude that Hlp is a human
representative of a conserved glucocorticoid-inducible cytochrome
P-450 gene family whose mechanism of induction involves
accumulation of Hlp mRNA.

FEATURES
source Location/Qualifiers
1. .2011

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/issue_type="liver"  
/clone_lib="human liver cDNA library provided from Meloy  
Laboratories"  
/notes="169 bp upstream of EcoRI site"  
61. .1575  
/codon_start=1  
/product="cytochrome P-450"  
/protein_id="BAA00001.1"  
/db_xref="GI:219962"  
/translation="MALIPDLAVETWLLAVSLVLLVYXGTHSHGLFKKLGIPGTPPL  
ZFLGNILSYHKGFCMFMECHCKYKGVYDQGPVLAITDPDMIKLVLVKECVSVF  
INREPFPGVGMKSAISAEDEEWKRLRSLSPFTSGKKNWPIIAQKGLVLRNL  
RRERETGKVTLDKDFGAYMDVITSSSFGVNDLNNPQDPLVENTKLLRDFLDLP  
FFLSITVFPFLIPLEVLNI CVFPRVNTFRKAVRMKESRLEDQKRVDFLOLMI  
DSHKNSKTESHSLDLVAQSIIFIFAGYETSSVLSFIMYSLATHPDVQOQLQE  
EIDAVLPKAPPTDVLQMEYLDWVNETLAFPIAMRLERVCKDKVEINGMFI PKG  
WVMVLPVALHRDPKYWTEPKELPERFSKKNNDIPYITPFGSGPRNCIGMRPAL  
NMKMLALIRVLQNSFPKCKETQIPLKLSLGLLQEPKPVLVKVESRDGTVSGA"
```

ORIGIN

Query Match	2.5%;	Score 48.4;	DB 9;	Length 2011;
Best Local Similarity	50.7%;	Pred. No. 0.012;		
Matches 146;	Conservative 0;	Mismatches 136;	Indels 6;	Gaps 1;
QY	997	AAAGCTGTCGAGGAGGAGTACTCGCTCAGATCAGTAACTGTTATTTGCTGGATATG	1056	
DB	926	AAGCTGTGCCATCTGAGCTCGTGGCCCAATCAATATTTATTTTGTGGCTATG	985	
QY	1057	AAACTTCTTCGACAGCTTTGACATGATGTTTCCCGACTCTCAGAGACAAAGCCGTTTC	1116	
DB	986	AAACCAGGACGAGTGTCTCTCTTCAATATGATGAATCGGCACCTCACCTGATGTC	1045	
QY	1117	AGGATAAAGCTTCGAGAGAAATTTGTCAGATCGACGAGATGCTTACCTAGCAGAAC	1176	
DB	1046	AGCAGAACTGCAGAGGAAATTTGATGAGTTTACCAATAGGCACCCACCTATG	1105	
QY	1177	TTAATGCGTTCG-----CTTATCTCGAAGCGTGTGTTAAGGAGTCTCTCGTCTAGACC	1230	
DB	1106	ATACGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC	1165	
QY	1231	CTCCTAGTCCGATGCTAACCGTGAATGCTTAAAGGATGAAGACTTCA	1278	
DB	1166	CAATTGCTATGAGACTTGAGAGGCTCTGCAAAAAGATGTTGAGATCA	1213	

Search completed: April 6, 2004, 16:19:22
Job time : 7671 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

'OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 08:05:34 ; Search time 833 Seconds

(without alignments)
9852.964 Million cell updates/sec

Title: US-10-066-007-2

Perfect score: 1932
Sequence: 1 gaattcgacagagccacc.....agcggctcgtcggaattc 1932

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

-Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1932	100.0	1932	3	AA64472 CDNA enco
2	353.2	18.3	3969	3	AA64473 DNA enco
3	51.8	2.7	2000	7	ADA71938 Rice gene
4	49.4	2.6	1501	6	ABL89830 Human pol
5	48.4	2.5	991	6	ABK97688 DNA enco
6	48.4	2.5	1062	6	ABK97694 DNA enco
7	48.4	2.5	1365	6	ABK97693 DNA enco
8	48.4	2.5	1512	2	AAQ87717 Human cyt
9	48.4	2.5	1512	2	AAT17399 Human der
10	48.4	2.5	1512	2	AAT28383 Human cyt
11	48.4	2.5	1512	6	ABK97692 DNA enco
12	48.4	2.5	1512	6	ACF06054 Human cyt
13	48.4	2.5	2011	6	ABN95088 Gene #158
14	48.4	2.5	2759	2	AAK28235 Human CYP
15	48.4	2.5	2764	7	ABX77172 DNA seque
16	48.4	2.5	2849	6	ABN95612 Gene #111
17	48.2	2.5	312	7	ABX39446 Bovine ES
18	48.2	2.5	1645	3	AAC77026 Human ORF
19	47.6	2.5	3059	3	AAC99040 Human pan
20	47.4	2.5	2160	2	AAZ10641 CDNA enco
21	47	2.4	445	5	ABV39053 Human pro
22	46.8	2.4	1712	2	AAV34249 Human sec
23	46.8	2.4	1712	7	ACD08120 CDNA enco

24	46.8	2.4	1822	2	AAV34248 Human sec
25	46.8	2.4	1822	7	ACD08119 CDNA enco
26	46.8	2.4	6004	9	ADB58866 Toxicity-
27	46.4	2.4	1552	4	ABL14451 Drosophil
28	45.4	2.3	576	5	ABV44839 Human pro
29	45.2	2.3	413	7	ABK97643 DNA enco
30	45.2	2.3	1608	6	ABK97643 DNA enco
31	45.2	2.3	1608	9	ADD29816 Human tum
32	45.2	2.3	1971	6	ABL62314 Colon ade
33	45.2	2.3	1971	6	ABL64874 Lung carc
34	45.2	2.3	1971	6	ABN95091 Gene #158
35	45.2	2.3	1971	6	ABN95091 Gene #158
36	45.2	2.3	1971	6	ABN95091 Gene #158
37	45.2	2.3	2010	3	AAK36214 Human cyt
38	45.2	2.3	2379	4	AAH79851 Human sec
39	45.2	2.3	2379	4	AAH57474 Human liv
40	45.2	2.3	2379	4	AAH57474 Human liv
41	45.2	2.3	2379	4	AAH57474 Human liv
42	45.2	2.3	2379	4	AAH57474 Human liv
43	45.2	2.3	2379	4	AAH57474 Human liv
44	45.2	2.3	2379	4	AAH57474 Human liv
45	45.2	2.3	2379	4	AAH57474 Human liv
46	45.2	2.3	2379	4	AAH57474 Human liv
47	45.2	2.3	2379	4	AAH57474 Human liv
48	45.2	2.3	2379	4	AAH57474 Human liv
49	45.2	2.3	2379	4	AAH57474 Human liv
50	45.2	2.3	2379	4	AAH57474 Human liv
51	45.2	2.3	2379	4	AAH57474 Human liv
52	45.2	2.3	2379	4	AAH57474 Human liv
53	45.2	2.3	2379	4	AAH57474 Human liv
54	45.2	2.3	2379	4	AAH57474 Human liv
55	45.2	2.3	2379	4	AAH57474 Human liv
56	45.2	2.3	2379	4	AAH57474 Human liv
57	45.2	2.3	2379	4	AAH57474 Human liv
58	45.2	2.3	2379	4	AAH57474 Human liv
59	45.2	2.3	2379	4	AAH57474 Human liv
60	45.2	2.3	2379	4	AAH57474 Human liv
61	45.2	2.3	2379	4	AAH57474 Human liv
62	45.2	2.3	2379	4	AAH57474 Human liv
63	45.2	2.3	2379	4	AAH57474 Human liv
64	45.2	2.3	2379	4	AAH57474 Human liv
65	45.2	2.3	2379	4	AAH57474 Human liv
66	45.2	2.3	2379	4	AAH57474 Human liv
67	45.2	2.3	2379	4	AAH57474 Human liv
68	45.2	2.3	2379	4	AAH57474 Human liv
69	45.2	2.3	2379	4	AAH57474 Human liv
70	45.2	2.3	2379	4	AAH57474 Human liv
71	45.2	2.3	2379	4	AAH57474 Human liv
72	45.2	2.3	2379	4	AAH57474 Human liv
73	45.2	2.3	2379	4	AAH57474 Human liv
74	45.2	2.3	2379	4	AAH57474 Human liv
75	45.2	2.3	2379	4	AAH57474 Human liv
76	45.2	2.3	2379	4	AAH57474 Human liv
77	45.2	2.3	2379	4	AAH57474 Human liv
78	45.2	2.3	2379	4	AAH57474 Human liv
79	45.2	2.3	2379	4	AAH57474 Human liv
80	45.2	2.3	2379	4	AAH57474 Human liv
81	45.2	2.3	2379	4	AAH57474 Human liv
82	45.2	2.3	2379	4	AAH57474 Human liv
83	45.2	2.3	2379	4	AAH57474 Human liv
84	45.2	2.3	2379	4	AAH57474 Human liv
85	45.2	2.3	2379	4	AAH57474 Human liv
86	45.2	2.3	2379	4	AAH57474 Human liv
87	45.2	2.3	2379	4	AAH57474 Human liv
88	45.2	2.3	2379	4	AAH57474 Human liv
89	45.2	2.3	2379	4	AAH57474 Human liv
90	45.2	2.3	2379	4	AAH57474 Human liv
91	45.2	2.3	2379	4	AAH57474 Human liv
92	45.2	2.3	2379	4	AAH57474 Human liv
93	45.2	2.3	2379	4	AAH57474 Human liv
94	45.2	2.3	2379	4	AAH57474 Human liv
95	45.2	2.3	2379	4	AAH57474 Human liv
96	45.2	2.3	2379	4	AAH57474 Human liv
97	45.2	2.3	2379	4	AAH57474 Human liv
98	45.2	2.3	2379	4	AAH57474 Human liv
99	45.2	2.3	2379	4	AAH57474 Human liv
100	45.2	2.3	2379	4	AAH57474 Human liv

ALIGNMENTS

RESULT 1
AAA64472
ID AAA64472 standard; CDNA; 1932 BP.
XX
AC AAA64472;
XX
DT 15-SEP-2003 (revised)
DT 02-JAN-2001 (first entry)
XX
DE CDNA encoding an astaxanthin synthetase polypeptide.
XX
KW Astaxanthin synthetase; astaxanthin; beta-carotene; carotenogenic yeast;
KW antioxidant; cancer; colouring reagent; farmed fish; salmon; ss.
XX
OS Xanthophyllomyces dendrorhous.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 33..1706
FT /*tag= a
FT /product= "astaxanthin synthetase"
FT polyA_signal
FT 1871
FT /*tag= b
XX
EP1035206-A1.
XX
PD 13-SEP-2000.
XX
PF 03-MAR-2000; 2000EP-00104430.
XX
PR 03-MAR-1999; 99EP-00104668.
PR 01-FEB-2000; 2000EP-00101666.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Hoshino T, Ojima K, Setoguchi Y;
XX
DR WPI; 2000-559874/52.
XX
P-PSDB; AAB08713.
XX
XX Novel polynucleotide encoding astaxanthin synthase useful for producing
XX recombinant cells for producing astaxanthin from beta-carotene.
XX Claim 3; Page 24-27; 46pp; English.
XX
CC The present sequence encodes an astaxanthin synthetase polypeptide of
CC Pfaffia rhodozyma. The enzyme is involved in the last step of the

CC astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.
 CC rhodospirillum rubrum yeast strain. The astaxanthin synthetase
 CC polynucleotides and polypeptides are useful for producing astaxanthin.
 CC Astaxanthin is an antioxidant which may be used to protect living cells
 CC against diseases such as cancer. Astaxanthin is also used as a coloring
 CC reagent, e.g. in farmed fish like salmon to impart an orange-red
 CC coloration. (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 1932 BP; 495 A; 456 C; 481 G; 500 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1932; DB 3; Length 1932;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATTCGGACGAGGCGACCTCTTCTCCATATGTTTCATCTTGGTCTTGTCTCAGGTG 60
 DB 1 GAATTCGGACGAGGCGACCTCTTCTCCATATGTTTCATCTTGGTCTTGTCTCAGGTG 60
 QY 61 CTTTAGGCTGGCTGCTTCTCATAGGATCCATAGCGTCTTCTCAGTCTTTACCTCGCTC 120
 DB 61 CTTTAGGCTGGCTGCTTCTCATAGGATCCATAGCGTCTTCTCAGTCTTTACCTCGCTC 120
 QY 121 CGAGGCGATCTTCACTGTATAACCTTCAGGCGCCGAAATCATACCAACTACTTTACAGGCA 180
 DB 121 CGAGGCGATCTTCACTGTATAACCTTCAGGCGCCGAAATCATACCAACTACTTTACAGGCA 180
 QY 181 ATTTTATAGACATCTCTCAGTCTGTACAGGTGAAGCATGCGAAGTACAGAGAAAT 240
 DB 181 ATTTTATAGACATCTCTCAGTCTGTACAGGTGAAGCATGCGAAGTACAGAGAAAT 240
 QY 241 ACGGAAGCACCCCTCGGTTTGTGGGATCGCTTGGACACCCGCTCTTGAACCTCGACCGATC 300
 DB 241 ACGGAAGCACCCCTCGGTTTGTGGGATCGCTTGGACACCCGCTCTTGAACCTCGACCGATC 300
 QY 301 CGAAGTCTTCAACATGTGATGAAGAGCTAGACTATCCGAACCTGGTATGCGCG 360
 DB 301 CGAAGTCTTCAACATGTGATGAAGAGCTAGACTATCCGAACCTGGTATGCGCG 360
 QY 361 CTCGAGTGTCTCAGAAATGCTACCGGAGATGTTGTTTACCGCGAAGGTGAAGCTCATA 420
 DB 361 CTCGAGTGTCTCAGAAATGCTACCGGAGATGTTGTTTACCGCGAAGGTGAAGCTCATA 420
 QY 421 ACGGACATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGTC 480
 DB 421 ACGGACATCGAAGGATCATGATCCCTCTCTGTCCGCTCAGGCGGTTAAGTCGATGTC 480
 QY 481 CAATTTCTTAGAAAAGGTATGGAACCTGTGCAAGATGATGGAGATCGCGCTCAGA 540
 DB 481 CAATTTCTTAGAAAAGGTATGGAACCTGTGCAAGATGATGGAGATCGCGCTCAGA 540
 QY 541 AGGATATGCGGTGGAGAGTCCGCGGTGAAAAGAGGCAACAGACTCGAGACCGAAG 600
 DB 541 AGGATATGCGGTGGAGAGTCCGCGGTGAAAAGAGGCAACAGACTCGAGACCGAAG 600
 QY 601 GAGTCCATGTAAGGATTCGGTCCGTCGAGTCTCTGAGCTCATGCTCTTCAGGAT 660
 DB 601 GAGTCCATGTAAGGATTCGGTCCGTCGAGTCTCTGAGCTCATGCTCTTCAGGAT 660
 QY 661 TTGACTATAAGAGCGACTCGCTCCAGAACAGACCAATGAGTCTATGTCTGCTTTGTCG 720
 DB 661 TTGACTATAAGAGCGACTCGCTCCAGAACAGACCAATGAGTCTATGTCTGCTTTGTCG 720
 QY 721 GACTTACCGATGGGTTGCTCTTCTGAGTCTGAGTCTCAAGGCTCATGTGGATTTTG 780
 DB 721 GACTTACCGATGGGTTGCTCTTCTGAGTCTGAGTCTCAAGGCTCATGTGGATTTTG 780
 QY 781 TACCTTACTTCGAACTATGAACCGGAGACATGAGTACCTTTTGACTCAAGGATTAGCAG 840
 DB 781 TACCTTACTTCGAACTATGAACCGGAGACATGAGTACCTTTTGACTCAAGGATTAGCAG 840
 QY 841 TTTCCCGACGAGTTGGATTCGAGCTTATGGACCAAAAGAGGCGCGTCTGGCTCAG 900
 DB 841 TTTCCCGACGAGTTGGATTCGAGCTTATGGACCAAAAGAGGCGCGTCTGGCTCAG 900

QY 901 CTTCCGATCAGCTGTTGTGATAAAAGGATGTTCAAGGTGCGGATATCTTAAGTCTCTAG 960
 DB 901 CTTCCGATCAGCTGTTGTGATAAAAGGATGTTCAAGGTGCGGATATCTTAAGTCTCTAG 960
 QY 961 TGAGAGCAAAATCATCGCGCAACCTGCTGAATCTCAAAAGCTGTCCGATGAGAGGTAC 1020
 DB 961 TGAGAGCAAAATCATCGCGCAACCTGCTGAATCTCAAAAGCTGTCCGATGAGAGGTAC 1020
 QY 1021 TCCTTCAGATCAGTAACTCTTATTGCTGATATGAAATCTTCTCGACAGTCTTGACAT 1080
 DB 1021 TCCTTCAGATCAGTAACTCTTATTGCTGATATGAAATCTTCTCGACAGTCTTGACAT 1080
 QY 1081 GGAATGTTTCCCGACTCTCAGAGACAAACCGCTCAGGATTAACCTTCAGAGAAATTT 1140
 DB 1081 GGAATGTTTCCCGACTCTCAGAGACAAACCGCTCAGGATTAACCTTCAGAGAAATTT 1140
 QY 1141 GTCAGATCGACACGGATATGCTTACGCTAGACGAACTTAATGCTTGCCTTATCTCGAAG 1200
 DB 1141 GTCAGATCGACACGGATATGCTTACGCTAGACGAACTTAATGCTTGCCTTATCTCGAAG 1200
 QY 1201 CGTTTGTAAAGAGTCTTCTGCTAGACCCCTGCTAGTCCGTATGCTAACCGTGAATGCT 1260
 DB 1201 CGTTTGTAAAGAGTCTTCTGCTAGACCCCTGCTAGTCCGTATGCTAACCGTGAATGCT 1260
 QY 1261 TAAAGGATGAAGACTTTCATCCCACTTCCGAGGCTGTCAATTGCTCGAGATGGTTCGCTCA 1320
 DB 1261 TAAAGGATGAAGACTTTCATCCCACTTCCGAGGCTGTCAATTGCTCGAGATGGTTCGCTCA 1320
 QY 1321 TCAAGAGGTCCGATCAGAGAAAGGATGCTGCTGCTTCCGTTGTTTCAACATCAATC 1380
 DB 1321 TCAAGAGGTCCGATCAGAGAAAGGATGCTGCTGCTTCCGTTGTTTCAACATCAATC 1380
 QY 1381 GTTCAAGTTCATTTATGAGAGATGTCAGAGAAATTCAGACCGGAGAGGTGCTTGAGG 1440
 DB 1381 GTTCAAGTTCATTTATGAGAGATGTCAGAGAAATTCAGACCGGAGAGGTGCTTGAGG 1440
 QY 1441 ACGTAAACAGATCTCGCTCAACAGATTTGAAGACCCCTATGACACCCAGGCGAGCTTATCT 1500
 DB 1441 ACGTAAACAGATCTCGCTCAACAGATTTGAAGACCCCTATGACACCCAGGCGAGCTTATCT 1500
 QY 1501 CTGGACCCAGAGCTTGTGTTGGTGGCGATTGCTGTCGCCGAGATGAAGGCTTCTTGT 1560
 DB 1501 CTGGACCCAGAGCTTGTGTTGGTGGCGATTGCTGTCGCCGAGATGAAGGCTTCTTGT 1560
 QY 1561 TTGTCACTCTCCGTCGGGTCCAGTTCGAGCCCATCATCTCATCCAGAGTACGAGCA 1620
 DB 1561 TTGTCACTCTCCGTCGGGTCCAGTTCGAGCCCATCATCTCATCCAGAGTACGAGCA 1620
 QY 1621 TCACCTTGATCATTTCCGCTCCTCGATCGTTGGTAGAGAGAGGAGGTTACCATGC 1680
 DB 1621 TCACCTTGATCATTTCCGCTCCTCGATCGTTGGTAGAGAGAGGAGGTTACCATGC 1680
 QY 1681 GTTTCAGGTCAAGCCCGCTCGAATGAGTTCATATGTTAAGAGAAAGTCTTATAT 1740
 DB 1681 GTTTCAGGTCAAGCCCGCTCGAATGAGTTCATATGTTAAGAGAAAGTCTTATAT 1740
 QY 1741 CTGAGATGTTGATGAGGACATGCTTCTTGTATCGATTTGTTCTCATACCCGGGC 1800
 DB 1741 CTGAGATGTTGATGAGGACATGCTTCTTGTATCGATTTGTTCTCATACCCGGGC 1800
 QY 1801 AGGCGCTATGACTTCTACGTCGTCTATGCTCGCTCTGAGTCTCTTCTTACCTATAT 1860
 DB 1801 AGGCGCTATGACTTCTACGTCGTCTATGCTCGCTCTGAGTCTCTTCTTACCTATAT 1860
 QY 1861 TATTCATCCGAAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 1920
 DB 1861 TATTCATCCGAAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 1920
 QY 1921 CGTCCGGAATTC 1932
 DB 1921 CGTCCGGAATTC 1932

FT	/tag= t
FT	/number= 9
FT	2236. .2354
FT	/tag= u
FT	/number= 10
FT	2355. .2431
FT	/tag= v
FT	/number= 10
FT	2432. .2542
FT	/tag= w
FT	/number= 11
FT	2543. .2618
FT	/tag= x
FT	/number= 11
FT	2619. .2652
FT	/tag= y
FT	/number= 12
FT	2653. .2742
FT	/tag= z
FT	/number= 13
FT	2963. .3050
FT	/tag= ac
FT	/number= 14
FT	3051. .3113
FT	/tag= ad
FT	/number= 14
FT	3114. .3171
FT	/tag= ae
FT	/number= 15
FT	3172. .3247
FT	/tag= af
FT	/number= 15
FT	3248. .3321
FT	/tag= ag
FT	/number= 16
FT	3322. .3398
FT	/tag= ah
FT	/number= 16
FT	3399. .3423
FT	/tag= ai
FT	/number= 17
FT	3424. .3513
FT	/tag= aj
FT	/number= 17
FT	3514. .3700
FT	/tag= ak
FT	/number= 18
FT	3865. .3866
FT	/tag= al
FT	
XX	
XX	BP1035206-A1.
PN	
XX	13-SEP-2000.
PD	
PP	
PF	
PP	03-WAR-2000; 2000EP-001044343
PR	
PR	09-WAR-1999; 93EP-001046666
PR	01-FEB-2000; 2000EP-001016666
XX	(HOFF) HOFFMANN LA ROCHE &
PA	
PA	Hoshino T, Ojima K, Setoguchi Y
PPI	
XX	WPI; 2000-559874/52.
DR	
DR	P-PADB; AAB08713.
XX	

PT Novel polynucleotide encoding astaxanthin synthase useful for producing
PT recombinant cells for producing astaxanthin from beta-carotene.

PS Claim 4; Page 30-33; 46pp; English.

XX The present sequence encodes an astaxanthin synthetase polypeptide of
XX Phafnia rhodomyza. The enzyme is involved in the last step of the
CC astaxanthin biosynthesis pathway, from beta-carotene to astaxanthin. P.
CC rhodomyza is a carotenogenic yeast strain. The astaxanthin synthetase
CC polynucleotides and polypeptides are useful for producing astaxanthin.
CC Astaxanthin is an antioxidant which may be used to protect living cells
CC against diseases such as cancer. Astaxanthin is also used as a colouring
CC reagent, e.g. in farmed fish like salmon to impart an orange-red
CC coloration. (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 3969 BP; 937 A; 942 C; 869 G; 1221 T; 0 U; 0 Other;

Query Match 18.3%; Score 353.2; DB 3; Length 3969;
Best Local Similarity 99.2%; Pred. No. 3.8e-87;
Matches 355; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1514 TTGCTTTGGTGGGATTTGCTGTCGCGAGATGAGGCTTCTTTGTCTACTCTCCG 1573
DB 3508 TCGCAGTGTGTGGGATTTGCTGTCGCGAGATGAGGCTTCTTTGTCTACTCTCCG 3567
QY 1574 TGGGGTCCAGTTCGAGCCCATCATCTCTCATCCAGATGACGACATCACCITGATCAT 1633
DB 3568 TGGGGTCCAGTTCGAGCCCATCATCTCTCATCCAGATGACGACATCACCITGATCAT 3627
QY 1634 TTCCCGTCTCGAATGCTTGTGTAGAGAGAGAGGGGTACAGATGCGTTTGCAGGTCAA 1693
DB 3628 TTCCCGTCTCGAATGCTTGTGTAGAGAGAGAGGGGTACAGATGCGTTTGCAGGTCAA 3687
QY 1694 GCGGTCGAATGAGTTGATTTCTCATATGTTAAGAGAGAGTTCTATCTGAGATGTG 1753
DB 3688 GCGGTCGAATGAGTTGATTTCTCATATGTTAAGAGAGAGTTCTATCTGAGATGTG 3747
QY 1754 ACTAGACAATGCTTTCTTTGTATCGATTTGTTCTCATACCGGCGAGCGCTATGACT 1813
DB 3748 ACTAGACAATGCTTTCTTTGTATCGATTTGTTCTCATACCGGCGAGCGCTATGACT 3807
QY 1814 TCTACGCTCTATGCTGCTGCTGACTCTCTTCTACCTATATATATTCATCCG 1871
DB 3808 TCTACGCTCTATGCTGCTGCTGACTCTCTTCTACCTATATATATTCATCCG 3865

RESULT 3

ADA71938
ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

PD 03-JAN-2003.

XX 22-JUN-2001; 2001WO-1B001105.

XX 22-JUN-2001; 2001WO-1B001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

XX Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 2.7%; Score 51.8; DB 7; Length 2000;
Best Local Similarity 9.8%; Pred. No. 0.0017;
Matches 72; Conservative 343; Mismatches 310; Indels 10; Gaps 3;

QY 421 AGGACATCGAAGGATCATGATCCCTCTCTGTCGCTCAGGCGGTTAAGTCGATCGTCC 480
DB 48 RKCSGCKNTTTRKKWYSASASGRTGSKWSSGYSKGMKKYKSKRWGRGRRG 107
QY 481 CAATTTTCTAGAAAAAGGTATGGAATTTCTGCAACAAGATGATGAGGATCGGCTCAGA 540
DB 108 MRSRMRWGRYRCARSGRAGSGRMMGKGRMSYMMCMYARCGGCKRKS--KGGS 165
QY 541 AGGATATGCGCTGGGAGAGTCCGCGGTGAAAAGAGGCAACACAGCTCGAGACGAAG 600
DB 166 MGKTCRRGARGSGWSGAKYKSGMSKRWMSGCGGRRSAYRYTGTSTKTYTK 225
QY 601 GAGTCGATGTAAGGATTTGGTTCGTCGACTACTCTGACGCTCATGCTTTTCAGGAT 660
DB 226 KWTYSASRCVRYMTTYSWACSSYTWCRSKRSMMWMMKRWMSRSYGMYSWKYM 285
QY 661 TTGACTATAAGGAGCATCGCTCCAGAAACAAGCAATGAGCTCTATGCTGCTTTGTGTCG 720
DB 286 MCTAYKKSYYSRWYRGSGWGRGATRYWGRGYMMAMMYKMYWYRGYKGMKRWAG 345
QY 721 GACTTACGATGCGTGTTCCTCTACCTGACGCTGCTCAAGGCTATCATGTGGGATTTG 780
DB 346 RMMRSMCRWSKACYWWRWWRWMTERRRWAKSSRTSRKKKCMKWKYKMRGYSR 405
QY 781 TACCTTACTTCGAACTATGAAACGAGACATGAGA-----TACCTTTGACTCAAGGAT 834
DB 406 MRSCKRARMKRCRSGRWKMCGRCMTCRMKSYGMWMSKRWMSKYSKMSRMRYWRK 465
QY 835 TAGCAGTTTCCGACGAGTTGGATCGAGCTTATGAGCAAAAAGAGCAGCGCTGCTG 894
DB 466 KCSSTTWTGTRGMMGTGRCYKRSKMKRCRRRWGRMYRWKRYNYSARYTWR 525
QY 895 GCTCAGTTTCCGATCAGCTTTGATAAAAAGGATGTTCAAGGTCGGGATATCTAAGTC 954
DB 526 YCARKKYSYSAARKARCWYRGKYWAGMMWKRYMYKMMWYKRYKYSKCSWYCKMS 585
QY 955 TCCTAGTGAGAGCAACATCGCC--GCCAACCTGCTGATCTCAAAAGCTGTCGATGA 1012
DB 586 YYASCWKSAKAGKMKRCKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMS 645
QY 1013 GGAGGTACTCGCTCAGATCAGTACCTGTTATTGCTGGATATGAAACCTCTTCGACAGT 1072
DB 646 SRWKSYYTCYWKMGSMKSTCTMYMSKYTYAKYGSYWRVYRWACWYMWYMYYSYRSY 705
QY 1073 CTTGACATGGATGTTTCCAGCACTCTCAGAGCAAAAGCGTTCAGGATTAACCTCAGA 1132
DB 1132 CTTGACATGGATGTTTCCAGCACTCTCAGAGCAAAAGCGTTCAGGATTAACCTCAGA 1132

706 MTYMAWYTSSTRNAMTKMYSGRYWTSMYKCYCKSCNKYSRSMYIYWSWNAKTMWMMRYA 765

1133 AGAAATTTCTCAGAT 1147

766 TRMMWYIRYSMKWY 780

RESULT 4

ABL89830

ID ID ABL89830 standard; cDNA; 1501 BP.

AC AC ABL89830;

XX XX

DT 24-MAY-2002 (first entry)

XX XX

DE Human polynucleotide SEQ ID NO 392.

XX XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KW anti-allergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein; gene; ss.

XX XX

OS Homo sapiens.

XX XX

WO200190304-A2.

XX XX

PD 29-NOV-2001.

XX XX

PF 18-MAY-2001; 2001WO-USO16450.

XX XX

PR 19-MAY-2000; 2000US-0205515P.

XX XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX XX

PI Birse CE, Rosen CA;

XX XX

DR WPI; 2002-122018/16.

XX XX

DR P-PSDB; ABB9421.

XX XX

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders.

PS Claim 4; SEQ ID NO 392; 2081pp + Sequence Listing; English.

XX XX

CC The invention relates to novel genes (ABL89449-ABL90853) and proteins

CC (ABB99040-ABB9044) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in

CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and

CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,

CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune

CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic

CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)

CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing

CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)

CC infectious diseases such as viral, bacterial, fungal and parasitic

CC infections. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX XX

SQ Sequence 1501 BP; 375 A; 389 C; 436 G; 296 T; 0 U; 5 Other;

Query Match 2.6%; Score 49.4; DB 6; Length 1501;

Best Local Similarity 63.2%; Pred. No. 0.007;

Matches 74; Conservative 1; Mismatches 42; Indels 0; Gaps 0

1785 TTCTCATACCGGGGAGCGGTATGATCTTACGCTGCTATCGCTCTGACTCTC 184

Db 522 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 581
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATGCTTAAGGATGAAGACTTCA 1278
Db 582 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAAGATGTTGAGATCA 629
RESULT 6
ABK97694
ID ABK97694 standard; DNA; 1062 BP.
XX
AC ABK97694;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA encoding cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 7-13.
XX
KW Cytochrome P450; CYP3A41; CYP3A42; CYP3A43; CYP3A4; CYP3A5; CYP3A7;
KW drug metabolism; drug design; drug screening; gene; ds.
XX
OS Homo sapiens.
XX
PN WC200244213-A1.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-SE002631.
XX
PR 28-NOV-2000; 2000SE-00004366.
PR 11-JUN-2001; 2001SE-00002061.
XX
PA (ZAPH/) ZAPHIROPOULOS P G.
PA (FINT/) FINTA C.
XX
PI Zaphiropoulos PG, Finta C;
XX
WPI; 2002-557532/59.
DR P-PSDB; ABG68755.
XX
Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
PT metabolism, in drug design and drug screening.
XX
PS Disclosure; Fig 13; 131pp; English.
XX
CC The invention describes a cytochrome P450 protein (I) in which CYP3A43
CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
CC fragments, variants and multiples of (I) having essentially the same
CC characteristics. (I) is useful as a medicament, and for evaluating drug
CC metabolism, in drug design, and drug screening, and in tests for
CC adjusting the dose of drugs. This sequence encodes a novel cytochrome
CC P450 protein
XX
SQ Sequence 1062 BP; 318 A; 239 C; 216 G; 289 T; 0 U; 0 Other;
Query Match 2.5%; Score 48.4; DB 6; Length 1062;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTATTCCTGATG 1056
Db 413 AAGCTCTGCTCCGATCTGGAGCTCGTGGCCCAATCAATATCTTTATTTTCTGGCTATG 472
QY 1057 AAACCTTCTTCGACGCTTTGACATGATGTTTCACCGACTCTCAGAAGACAAGCCGTTTC 1116
Db 473 AAACCAAGGACGATGTTCTCTCTTCATATGATGTAAGTGGCCACTCACCTGATGTC 532
QY 1117 AGGATAAATTCGAGAAGAAATTTGTCAGATCGACACGGATATGCTACGCTAGACGAAC 1176
Db 533 AGCAGAAATCTGCAGGAGGAATATGATGCAAGTTTACCCCAATAAGGCACCAACCTATG 592
QY 1177 TTAATGGTTC-----CTTATCTCGAAGCGTTTGAAGGAGTCTCTTCGCTAGACC 1230

Db 593 ATACTGTGCTACAGATGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 652
QY 1231 CTCCTAGTCCGATGCTAACCGTGAATGCTTAAGGATGAAGACTTCA 1278
Db 653 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAAGATGTTGAGATCA 700
RESULT 7
ABK97693
ID ABK97693 standard; DNA; 1365 BP.
XX
AC ABK97693;
XX
DT 07-OCT-2002 (first entry)
XX
DE DNA encoding cytochrome P450 CYP3A43 exon 1-CYP3A4 exon 4-13.
XX
KW Cytochrome P450; CYP3A41; CYP3A42; CYP3A43; CYP3A4; CYP3A5; CYP3A7;
KW drug metabolism; drug design; drug screening; gene; ds.
XX
OS Homo sapiens.
XX
PN WC200244213-A1.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-SE002631.
XX
PR 28-NOV-2000; 2000SE-00004366.
PR 11-JUN-2001; 2001SE-00002061.
XX
PA (ZAPH/) ZAPHIROPOULOS P G.
PA (FINT/) FINTA C.
XX
PI Zaphiropoulos PG, Finta C;
XX
WPI; 2002-557532/59.
DR P-PSDB; ABG68754.
XX
Novel cytochrome P450 protein in which CYP3A43 exon 1 is joined to sets
PT of CYP3A4 or CYP3A5 exons, useful as medicament, and in evaluating drug
PT metabolism, in drug design and drug screening.
XX
PS Claim 9; Fig 12; 131pp; English.
XX
CC The invention describes a cytochrome P450 protein (I) in which CYP3A43
CC exon 1 is joined to sets of CYP3A4 or CYP3A5 exons, as well as sub
CC fragments, variants and multiples of (I) having essentially the same
CC characteristics. (I) is useful as a medicament, and for evaluating drug
CC metabolism, in drug design, and drug screening, and in tests for
CC adjusting the dose of drugs. This sequence encodes a novel cytochrome
CC P450 protein
XX
SQ Sequence 1365 BP; 406 A; 299 C; 294 G; 366 T; 0 U; 0 Other;
Query Match 2.5%; Score 48.4; DB 6; Length 1365;
Best Local Similarity 50.7%; Pred. No. 0.013;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
QY 997 AAAGCTGTCGATGAGGAGTACTCGCTCAGATCAGTAACTGTTATTCCTGATG 1056
Db 716 AAGCTCTGCTCCGATCTGGAGCTCGTGGCCCAATCAATATCTTTATTTTCTGGCTATG 775
QY 1057 AAACCTTCTTCGACGCTTTGACATGATGTTTCACCGACTCTCAGAAGACAAGCCGTTTC 1116
Db 776 AAACCAAGGACGATGTTCTCTCTTCATATGATGTAAGTGGCCACTCACCTGATGTC 835
QY 1117 AGGATAAATTCGAGAAGAAATTTGTCAGATCGACACGGATATGCTACGCTAGACGAAC 1176
Db 836 AGCAGAAATCTGCAGGAGGAATATGATGCAAGTTTACCCCAATAAGGCACCAACCTATG 895
QY 1177 TTAATGGTTC-----CTTATCTCGAAGCGTTTGAAGGAGTCTCTTCGCTAGACC 1230

Db 896 ATACTGTCTCAGATGAGTATCTTGACATGGTGGTGAATGAACCGCTCAGATTATTC 955
 QY 1231 CTCCTAGTCGGTATCCTAACCGTGATCTTAAGGATGAAGACTTCA 1278
 Db 956 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGAGATGTTGATCA 1003

RESULT 8
 AAQ87717
 ID AAQ87717 standard; cDNA; 1512 BP.
 XX
 AC AAQ87717;
 DT 25-MAR-2003 (revised)
 DT 14-NOV-1995 (first entry)
 XX
 DE Human cytochrome P450 molecular species 3A4 cDNA coding region.
 XX
 KW Human cytochrome P450; amplification; PCR; primer; expression vector;
 KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
 KW carcinogen; mutagen; liver metabolism; ds.
 XX
 OS Homo sapiens.
 XX
 PN EF644267-A2.
 XX
 PD 22-MAR-1995.
 XX
 PF 20-JUL-1994; 94EP-00111298.
 XX
 PR 20-JUL-1993; 93JP-00201120.
 PR 21-JUL-1993; 93JP-00180246.
 PR 30-JUL-1993; 93JP-00208279.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 PI Hayaishi K, Sakaki T, Yabueaki Y, Komai K, Kaneko H, Nakatsuka I;
 DR WPI; 1995-116991/16.
 DR P-PSDB; AAR72363.
 XX
 PT Evaluation of safety of a chemical cpd. - using recombinant yeast
 PT expressing human cytochrome P450 and a yeast NADPH-P450 reductase.
 XX
 PS Example; Page 31-33; 124pp; English.
 XX
 CC The nucleotide sequence of the cDNA coding region for the human
 CC cytochrome P450 species 3A4. The gene encodes a protein of 503 amino
 CC acids. The cDNA was amplified by PCR using the primers AAQ87743-6. The
 CC product was cloned into the yeast expression vectors pAAH5N or pAHRH to
 CC produce the vectors p3A4 for the expression of the cytochrome P450 alone
 CC or p3A4R for co-expression with the yeast NADPH-P450 reductase. The
 CC vectors are used in a method for evaluating the safety of a chemical
 CC compound by reacting the chemical compound with recombinantly produced
 CC human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715),
 CC 2B1 (AAQ87716), or 3A4 or their auxiliary species and variants (AAQ87718-
 CC 32), and yeast NADPH-P450 reductase, either as a fused protein or in cell
 CC extracts, and analysing the resulting metabolite to assess the safety of
 CC the chemical compound. The method is useful for determining whether the
 CC chemical compound, or its metabolite, will be converted into a
 CC carcinogenic or mutagenic form through metabolism in the liver. (Updated
 CC on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 1512 BP; 441 A; 330 C; 330 G; 411 T; 0 U; 0 Other;
 Query Match 2.5%; Score 48.4; DB 2; Length 1512;
 Best Local Similarity 50.7%; Pred. No. 0.013;
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
 QY 997 AAAAGCTGTCGATGAGGAGTACTCCCTCAGATCAGTAACCTGTTATTTCTGGATATG 1056
 Db 863 AAGCTCTGTCGATCTGGAGCTCGTGGCCCAATCAATATCTTTATTTTCTGCTGATG 922

QY 1057 AAATCTTTCGACAGTCTTGACATGGATGTTTACCGACTCTCAGAAGACAAGCCGTTTC 1116
 Db 923 AAACACGACGAGTGTCTCTCTCTCATATGATGATGACCTGCGCCACTCACCCTGATGTC 982
 QY 1117 AGGATAAATCTCGAGAGAGAAATTTGTTCAGATCGACACGGATATGCTAGCTAGCAGAAC 1176
 Db 983 AGCAGAAATCTCGAGGAGGAAATTTGATGACGTTTACCCCAATAGGCACCACCCTATG 1042
 QY 1177 TTAATGGTGTGC-----CTTATCTCGAAGCGTTTGTTAAGGAGTCTCTTCGTTCTAGACC 1230
 Db 1043 ATACTGTCTACAGATGGAGTATCTTGACATGGTGGTGAATGAACGCTCAGATTATTC 1102
 QY 1231 CTCCTAGTCCGTATGCTAACCGTGATCTTAAGGATGAAGACTTCA 1278
 Db 1103 CAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAGAGATGTTGATCA 1150

RESULT 9
 AAT17399
 ID AAT17399 standard; cDNA; 1512 BP.
 XX
 AC AAT17399;
 XX
 DT 01-AUG-1996 (first entry)
 XX
 DE Human derived cytochrome P4503A4 cDNA.
 XX
 KW Human derived cytochrome; P4503A4; commercial cDNA library; yeast;
 KW transfection; recombinant production; expression vector; mammal;
 KW immunisation; sensitisation; antibody; determination; detection;
 KW non-cross reactive; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1512
 FT /*tag= a
 XX
 PN JP08027197-A.
 XX
 PD 30-JAN-1996.
 XX
 PF 13-JUL-1994; 94JP-00161552.
 XX
 PR 13-JUL-1994; 94JP-00161552.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 WP1; 1996-136338/14.
 DR P-PSDB; AAR81464.
 XX
 PT Antibody recognising human derived cytochrome P4502A4 - allows specific
 PT detection of cytochrome P450 species in humans.
 XX
 PS Example 1; Page 10-12; 13pp; Japanese.
 XX
 CC The present sequence encodes the human derived cytochrome (HDC) P4503A4,
 CC which was obt'd. from a commercial cDNA library. Yeast were transfected
 CC with an expression vector contg. the HDC cDNA, cultured and then
 CC disrupted to give a microsomal fraction. The HDC was purified from the
 CC fraction, and used to immunise and sensitise a mammal. Blood was drawn
 CC from the mammal, and an anti-HDC antibody isolated. The antibody obt'd.
 CC recognises HDC P4503A4, partic. at a serum dilution rate of 1:10000, and
 CC is substantially without cross reaction to other HDC P450 spp
 XX
 SQ Sequence 1512 BP; 441 A; 330 C; 330 G; 411 T; 0 U; 0 Other;
 Query Match 2.5%; Score 48.4; DB 2; Length 1512;
 Best Local Similarity 50.7%; Pred. No. 0.013;
 Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
 QY 997 AAAAGCTGTCGATGAGGAGTACTCCCTCAGATCAGTAACCTGTTATTTCTGGATATG 1056

Example 1; Page 24-26; 74pp; Japanese.

This is the nucleotide sequence of the human cytochrome P450 molecular species 3A4 gene which encodes a protein of 503 amino acids. The gene was amplified from a human liver derived cDNA library as 2 fragments of 0.6 and 0.9 kb using primers AAR26933-6. The prod. was cloned into the yeast expression vector pAAH5N to generate plasmid p3A4 for prodn. of the cytochrome only or into the vector pAHR5 to generate the plasmid p3A4R for co-prodn. with the yeast NADPH-P450 reductase. The sequence is placed under control of the yeast ADH gene promoter and terminator. The vectors are used in a method for evaluating the safety of a cpd by reacting the test cpd. with recombinantly produced human cytochrome P450 mol. species 1A2 (AAR28380), 2C9 (AAR28381), 2B1 (AAR28382), 3A4 or their variants (AAR28384-98) together with yeast NADPH-P450 reductase (either as a fused protein or as a cell extract) and analysing the resultant metabolite. The cpd. is considered "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it is not detoxified or is metabolised to a carcinogenic cpd

CC P450 protein adjusting in

PA (GENE-) GENE STREAM PTY LTD.
XX FI Daly JM;
XX DR WPI; 2003-093021/08.
XX DR P-PSDB; ABU57260.
XX PT New transgenic non-human animal expressing a foreign polypeptide
PT associated with drug behavior and/or metabolism, useful for studying the
PT behavior and/or metabolism of a drug in other animals.
XX PS Example 2A; Page 296-299; 408pp; English.
XX CC This invention relates to a transgenic non-human animal which may be used
CC for assessing the behaviour and/or metabolism of a drug in another animal
CC and which expresses a foreign polypeptide associated with drug behaviour
CC and/or metabolism. The invention also comprises a nucleic acid construct
CC for use in producing the above transgenic non-human animal and a method
CC of assessing the metabolism and/or behavior of a drug in an animal of
CC interest, comprising administering a test agent to the transgenic animal
CC and conducting analytical tests to determine drug metabolism and/or
CC behaviour. The transgenic animal is useful in studying drug metabolism
CC and/or behaviour in other animals. The nucleic acid construct is useful
CC in producing the above transgenic animal and the methods are used for
CC producing, breeding and using transgenic animals for pharmacological
CC (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological
CC studies. Nucleic acid sequences used within the invention are serum
CC albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine
CC diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins
CC and (MRP's). The present sequence represents a DNA sequence used to
CC create a transgenic animal within the scope of the invention
XX SQ Sequence 2764 BP; 823 A; 588 C; 771 T; 0 U; 0 Other;
Query Match 2.5%; Score 48.4; DB 7; Length 2764;
Best Local Similarity 50.7%; Pred No. 0.018;
Matches 146; Conservative 0; Mismatches 136; Indels 6; Gaps 1;
QY 997 AAAAGCTGTCGATCAGGAGGTACTCGCTCAGATCAGTAACTCTTTATTTGCTGGATATG 1056
Db |||||
967 AAGCTCTGTCGATCAGGAGGTACTCGCTCAGATCAGTAACTCTTTATTTGCTGGCTATG 1026
QY 1057 AACTTCTTCGACATCTTGATCGATGTTTCCCGACTCTCAGAGACAAAGCCGTTTC 1116
Db |||||
1027 AACCCAGAGCAGTGTCTCTCTTCATTATGATGAACTGGCCACTCACCCTGATGCC 1086
QY 1117 AGGATAAACTTCGAGAAGAAATTTCTCAGATCGACACGGATATGCCTACGCTAGACGAAC 1176
Db |||||
1087 AGCAGAACTCGCAGGAGGAAATTTGATGCGAGTTTACCCTAATAAGGCACCCACCTATG 1146
QY 1177 TTAATCGGTGC-----CTTATCTCGAAGCGTTTGTAAAGAGTCTCTTCGTGACCC 1230
Db |||||
1147 ATACTGTGTACAGATGGAGTATCTTTCATGTCATGCTGGTGAATGAACCGCTCAGATTATTC 1206
QY 1231 CTCCTAGTCGATGCTTAACCGTGAATCTTAAAGGATGAAGACTTCA 1278
Db |||||
1207 CAATGCTATGAGACTTGAGAGGGTCTGCAAAAAGATGTTGAGATCA 1254